

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 27, 2001, 16:54:05 ; Search time 7648.39 Seconds

(Without alignments)  
5996.315 Million cell updates/sec

Title: US-09-830-647-3

Perfect score: 2780  
1 aatttcgcgcgcgcgcctctctg.....aaaaaaaaaaaaaactcgcag 2780

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : GenBank1:  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_ov:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_om:\*  
20: em\_or:\*  
21: em\_ov:\*  
22: em\_pat:\*  
23: em\_ph:\*  
24: em\_pl:\*  
25: em\_ro:\*  
26: em\_sts:\*  
27: em\_sy:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_htgo\_hum:\*  
31: em\_htgo\_inv:\*  
32: em\_htgo\_rod:\*  
33: em\_htg\_hum:\*  
34: em\_htg\_inv:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | DB | ID       | Description        |
|------------|--------|-------------|--------|----|----------|--------------------|
| 1          | 2780   | 100.0       | 2780   | 9  | AB028069 | AB028069 Homo sapi |
| 2          | 2648   | 95.3        | 2719   | 9  | AB028070 | AB028070 Homo sapi |
| 3          | 2458.8 | 88.4        | 2474   | 9  | AF160249 | AF160249 Homo sapi |
| 4          | 2456   | 88.3        | 2460   | 9  | AF160876 | AF160876 Homo sapi |
| 5          | 2118.6 | 76.2        | 236236 | 2  | AC018977 | AC018977 Homo sapi |
| 6          | 1537   | 55.3        | 2458   | 10 | AF292400 | AF292400 Cricetulu |
| 7          | 1236   | 44.5        | 2309   | 10 | MMDBF4   | MMDBF4 Mus muscu   |
| 8          | 1194.2 | 43.0        | 123331 | 9  | AC005164 | AC005164 Homo sapi |
| 9          | 1194.2 | 43.0        | 163031 | 2  | AC024947 | AC024947 Homo sapi |
| 10         | 1194.2 | 43.0        | 163889 | 2  | AC004958 | AC004958 Homo sapi |
| 11         | 1101.6 | 39.6        | 2276   | 9  | AK022969 | AK022969 Homo sapi |
| 12         | 1003.8 | 36.1        | 232903 | 2  | AC001341 | AC001341 Rattus no |
| 13         | 554.6  | 19.9        | 136823 | 2  | AC003083 | AC003083 Homo sapi |
| 14         | 513.8  | 18.5        | 685    | 6  | AX070167 | AX070167 Sequence  |
| 15         | 423.6  | 15.2        | 177899 | 2  | AC074175 | AC074175 Mus muscu |
| 16         | 414.4  | 14.9        | 433    | 11 | G36582   | G36582 SHGC-53468  |
| 17         | 411    | 14.8        | 422    | 6  | AX072042 | AX072042 Sequence  |
| 18         | 384.4  | 13.8        | 403    | 6  | AX070196 | AX070196 Sequence  |
| 19         | 332.6  | 12.0        | 205662 | 9  | AC017002 | AC017002 Homo sapi |
| 20         | 327    | 11.8        | 136555 | 2  | AC068279 | AC068279 Homo sapi |
| 21         | 327    | 11.8        | 157063 | 2  | AC083899 | AC083899 Homo sapi |
| 22         | 217.4  | 7.8         | 220    | 11 | G41942   | G41942 SHGC-64833  |
| 23         | 211.4  | 7.6         | 397    | 6  | AX070116 | AX070116 Sequence  |
| 24         | 126.2  | 4.5         | 168717 | 2  | AC024683 | AC024683 Homo sapi |
| 25         | 126.2  | 4.5         | 177720 | 9  | AC007514 | AC007514 Homo sapi |
| 26         | 104.6  | 3.8         | 1517   | 9  | AF125531 | AF125531 Homo sapi |
| 27         | 99.4   | 3.6         | 377    | 6  | AX070662 | AX070662 Sequence  |
| 28         | 67.8   | 2.4         | 1141   | 6  | AX083744 | AX083744 Sequence  |
| 29         | 65.6   | 2.4         | 1141   | 6  | AX083744 | AX083744 Sequence  |
| 30         | 64.2   | 2.3         | 87256  | 2  | AC013754 | AC013754 Homo sapi |
| 31         | 63.6   | 2.3         | 207420 | 2  | AC078884 | AC078884 Mus muscu |
| 32         | 63.4   | 2.3         | 321003 | 2  | PFMAL4P3 | PFMAL4P3 Plasmodu  |
| 33         | 62     | 2.2         | 7218   | 6  | I66494   | I66494 Sequence 14 |
| 34         | 61.2   | 2.2         | 71559  | 2  | AC091013 | AC091013 Homo sapi |
| 35         | 61     | 2.2         | 245802 | 2  | AC006279 | AC006279 Plasmodu  |
| 36         | 60.8   | 2.2         | 51274  | 2  | AC023454 | AC023454 Homo sapi |
| 37         | 60.8   | 2.2         | 167968 | 2  | AC084398 | AC084398 Homo sapi |
| 38         | 60.6   | 2.2         | 140176 | 2  | AC016411 | AC016411 Homo sapi |
| 39         | 60.6   | 2.2         | 174133 | 8  | AF083031 | AF083031 Guillard  |
| 40         | 60.2   | 2.2         | 201676 | 2  | AC090647 | AC090647 Mus muscu |
| 41         | 58.8   | 2.2         | 120091 | 2  | AC012684 | AC012684 Homo sapi |
| 42         | 59.8   | 2.2         | 170102 | 9  | AC008079 | AC008079 Homo sapi |
| 43         | 59.6   | 2.1         | 101509 | 2  | AC027353 | AC027353 Homo sapi |
| 44         | 59.4   | 2.1         | 62164  | 2  | AC068512 | AC068512 Homo sapi |
| 45         | 59.4   | 2.1         | 158167 | 2  | AC034105 | AC034105 Homo sapi |

## ALIGNMENTS

| RESULT     | 1          | AB028069   | 2780 bp | mrna | PRI | 03-JUN-1999 |
|------------|------------|--|---------|------|-----|-------------|
| LOCUS      | AB028069   | Homo sapiens mRNA for activator of S phase kinase, complete cds.   |         |      |     |             |
| DEFINITION | AB028069   | Homo sapiens mRNA for activator of S phase kinase, complete cds.   |         |      |     |             |
| ACCESSION  | AB028069.1 | GI:4996095   |         |      |     |             |
| VERSION    | AB028069.1 | GI:4996095   |         |      |     |             |
| KEYWORDS   |            | activator of S phase kinase.   |         |      |     |             |
| SOURCE     |            | Homo sapiens Hela cDNA clone:H37.  |         |      |     |             |
| ORGANISM   |            | Homo sapiens   |         |      |     |             |
| REFERENCE  |            | 1 (sites)  |         |      |     |             |
| AUTHORS    |            | Kumagai, H., Sato, N., Yamada, M., Mahony, D., Seghezzi, W., Lees, E.,   |         |      |     |             |
| TITLE      |            | A Novel Growth- and Cell Cycle-Regulated Protein, ASK, Activates Human Cdc7-Related Kinase and Is Essential for G1/S Transition in Mammalian Cells |         |      |     |             |
| JOURNAL    |            | Mol. Cell. Biol. (1999) In press   |         |      |     |             |
| REFERENCE  |            | 2 (bases 1 to 2780)  |         |      |     |             |
| AUTHORS    |            | Masai, H., and Kumagai, H.   |         |      |     |             |

| TITLE   | Direct Submission   |
|---------|---|
| JOURNAL | Submitted (28-MAY-1999) to the DBU/EMBL/Genbank databases. Hisao Msaai, Institute of Medical Science, University of Tokyo, Department of Developmental Biology: 4-6-1 Shirokanebashi, Minato-ku, Tokyo 108-8639, Japan (E-mail: hisao@ims.u-tokyo.ac.jp, Tel: 01-3-5449-5661, Fax: 01-3-5449-5424). |

**FEATURES**  
**SOURCE**

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="HeLa"
/chromosome="7"
/clone="H3"
/map="7q21"
518..2542
/feature="ASK/H37 (hubbfa)"
518..2542
gene
CDS

```

sequences.<sup>10</sup> The "functional binds and activates hucd7 kinase" domain (residues 101-120 and 552-564 are two putative *pestr*-like

```

/codon_start=1
/product="activator of S phase kinase"
/protein_id="BA078326.1"
/db_xref="gi:4996096"
/translation="MNSGAMRIHSGKHPGGIOVKNENRPSLKLKTNPPREKSKCIR
PLNGKFLYDLPSVNISEKLOKIDKLGGRVEEFLSKVSLINKEKKNFQNTGRI
SPVSPESAVYETTPSPHSDSGSPSDYLSRGVLVKAIRKHDPISNLSL
NALSKGKYLTIHDDIRYITDOKKRELILKSTSVRGCGKRGACGTRGSLKRL
FVYVPMISQLYRPFYLOLTNNPRLNYSIQPCSPDVPBMSNGKQVQYLIQYDGS
KYGSTIQQLKEKKKKCYGCECLQKEDYEDLHLLSEPHNFAVSNOIQVYDILSK
VEFVEIKDIPFKKRIKYSGLSPSVASVYLKIDOKKEVLOHISQKCEQDITIT
KEONFLKETOETELKFLISEPILPHSNELGKIKNSVNSQFSDINDSGSQPKR
DYLPHKROECLIDISEHTLASEDELDYRDKIQKASVSDSFSDINDSGSQPKR
DYLPAKDKOKEDSLIFTHDSGLTINSESHLYVAKPFTPEEPEENEDCNKRL
DLSLPSKIRHAKVILILGNKRENEPNAEPLARTEFYLQENRACSSPVASLIDLFQ
SEKSESLGFTSYTSGSCICNVLDIWEENSNDLLTAFFSSPTSTFTG"
.
2724 . 2730

```

| Query Match           | 100.0%         | Score 2780  | DB 9     | Length 2780 |
|-----------------------|----------------|---|----------|-------------|
| Best Local Similarity | 100.0%         | Pred. No 0  |          |             |
| Matches 2780          | Conservative 0 | Mismatches 0  | Indels 0 | Gaps 0      |
| QY                    | 1              | aattgcgaacgagctctctcgtgagctgagccaaagacttgaagcgcgagacagagcccg    | 60       |             |
| Db                    | 1              | AATTGGGACACAGGCTCTCTTAGGCTGCGCAAGACTGTAAGCGCGAGCCGAGACCCGG      | 60       |             |
| QY                    | 61             | gltctagactctagagagcaacggaatbgaagcgcggtagagcgagaaacaacactcgag    | 120      |             |
| Db                    | 61             | GTCTAGAGCTAGAGACCAACGGAAATGAGGCGGGGGGTATGAGCGCGAAACACAACTCGAG   | 120      |             |
| QY                    | 121            | ggccgaagcagagcgcgcgagaagaacgcgcgctgaagggcgcgagcgcgcaacgcgagaa   | 180      |             |
| Db                    | 121            | GGCGACAGACGAGCGCGCGAGAAAGACCGCGCGCTGAGGGGGCGCGCCACGCGGAA        | 180      |             |
| QY                    | 181            | ggcagaagcagagggcgcgagacggaagggcgagacgagcgctgagcgtgaaacagggcgagg | 240      |             |
| Db                    | 181            | GGCAGACACAGAGGGCGACGCGCAGCGCGGCGACGGCGCTGGCGTGAACCGGCGGGG       | 240      |             |
| QY                    | 241            | cgcgcgatctggcgcgcgcgcgcgctgaacggtttcaaatcttcaacgcgcgaagccca     | 300      |             |
| Db                    | 241            | CGCGGATCTGGCGCGCGCGCGCGCGCGAGACGCTTTTCAAACTTCACACGCGACGCCA      | 300      |             |
| QY                    | 301            | ctcgcttgagctttgcgctctcctcctcgcgcttgaagcggatccgagcccgaggaac      | 360      |             |
| Db                    | 301            | CTCGCTTGAGCTTTGCGCTCTCCTCCTCGCGCTTGAAGCGGATCCGAGCCCGAGGAAC      | 360      |             |

|    |      |  |      |
|----|------|--|------|
| Dh | 301  | CGTTTTGTCCTTGCGCTTCCTCCCGCGCGCTTGAGAGCGGATCCGGACCCCGGAAAC        | 360  |
| OY | 361  | ccgaactcgaacccgggtaactctactcgtgttagagaccgttagctcgtcgcggaagaagaag | 420  |
| Dh | 361  | CCGACCTCGAGACGGCGTACTCTACTCGCTAGAGCCCTACCTCGCGGAAGCAGAGAG        | 420  |
| OY | 421  | ggagcgtctctgtcaacagcggcgggggaagccgtgtcttcgagcgtctgcgggttcgagcat  | 480  |
| Dh | 421  | CGGCGGCTCTGTCAACAGGACGGAGGAGACCGTGTCTTCGCGGCTCCCGCTCGACAT        | 480  |
| OY | 481  | ttcttcggaaccgaacgagtgtaagtgccggcgagctgcattgaatccggagccatgaagat   | 540  |
| Dh | 481  | TTCTCCGAGCCCAAGCATATAGTGTCCGGGGAGCTGCAATGACCTCCGAGCAATGAGGAT     | 540  |
| OY | 541  | ccacagtaagaagacattccagggctggaatccaaagtcgaaaaaagaaaaacagacatc     | 600  |
| Dh | 541  | CCACAGTAAAGAGCATTTTCCAGGGTGGAAATCCAAAGTCAAAAATGAAAAAACAGACATTC   | 600  |
| OY | 601  | tcggaactcctcggaaactctgaatacagccgagaaataaccaaagtcacatcttcgggg     | 660  |
| Dh | 601  | TCGGAATCTCTGAAAACTGTATTAACAGGCCAGAAAAATCCAAAGTAAAGCACTTTGGGG     | 660  |
| OY | 661  | aaaagatattcttccttgacttaacctctgtgcacatatctgaanaactcnaaagaagat     | 720  |
| Dh | 661  | AAAAGATTATTTCCTTGACTTACTCTGTGCACATCTCGAAAACTTCAAAAGGACAT         | 720  |
| OY | 721  | taagagctcgggaaggcgagttgaagaatttcgagaaagatactagttatctcttc         | 780  |
| Dh | 721  | TAAAGATCTGGAGGGGAGTGAAGATTTCAGCAAAAGATATCAGTTATCTTATTTTC         | 780  |
| OY | 781  | aaataagaagagaagtaatttgacaaacacttggatgcgaattctctctgtaccagatcc     | 840  |
| Dh | 781  | AATTAAGACGAGAGCTTAATTGTGCACAAACCTTGGGTGCAATTTCTCTGTACCAAGTCC     | 840  |
| OY | 841  | agaaatcgatattactggaagaacacacttaactatccacggcgagatgaagaattcat      | 900  |
| Dh | 841  | AGAAATCGATATTACTGCAAGAACCACTTCACTCATCCAGCCATATGGAAGTTCAAT        | 900  |
| OY | 901  | taagtaaccagacagacgtgtgtttaagcagagggaaaaattatgtatgaanaagctatcaa   | 960  |
| Dh | 901  | TAAGTACCAAGACAGAGTGTTTAAACAGAGAAATATATAGTGAAMAAACCTATCA          | 960  |
| OY | 961  | ggaccatgattctatctccctcaaatagtatataccaatgcctctgtcatcggggagtaaa    | 1020 |
| Dh | 961  | GGACCATGATTTTATTTCTCTTCAAAATAGTATTTATTCAAATGCTTGTCAATGGGAGATAA   | 1020 |
| OY | 1021 | aattcttcatalitgtagcatllagatlaacatctgaacaaagaanaaagaagcttgcattc   | 1080 |
| Dh | 1021 | AATTTCTATTTATGTATGATGACATTAATAATCTAATGTACAAAGAAAAAGAAAGTTGATTT   | 1080 |
| OY | 1081 | acccaagaatcaatgaattactcttgaagaagatbgygggcanaaagaatttgtagtggtagaa | 1140 |
| Dh | 1081 | ACTCAAGAATCATAGTACTTCTAGTAAGACGATGGGGCAAAAAGATTTGGTAGTGTGCACA    | 1140 |
| OY | 1141 | aaaaaacaagaacagaanaaacccaanaagcctcttbaaaggtggaagatactgagccaact   | 1200 |
| Dh | 1141 | AAAAACAAACAGAGAACATCACTCAAAAAGCCTTTGTGAAGAAGTGAMACATGTACGCCAAT   | 1200 |
| OY | 1201 | ttataggaatttctcttgaagctggagccaatactgcctcttataatattctatccaaga     | 1260 |
| Dh | 1201 | TTATAGGCATTTTATCTTCAAGCTGACGACCAATATGCGCTTTTAAATTTATTTCTATTGAGA  | 1260 |
| OY | 1261 | ggccctcagtcattatgtatgagaacagcaatctagtatggaagaagaacacttaagttaa    | 1320 |
| Dh | 1261 | GGCCTCAAGTCTATTTGATGTATGTAGAACGCCATCTAGTATGCMAAGCAACACTGAGTTAA   | 1320 |
| OY | 1321 | actaagaatccaacacagatggcgaatgaatgagtggaacctcaatcaactcagttgaa      | 1380 |
| Dh | 1321 | ACTTAAGAATTCAAACAGATGGCGATTAAGTATGTGTGGAACCTCAATTCACAGTTGAA      | 1380 |
| OY | 1381 | aggaagaagaaaaaagataattgtgaaattgtgccttgtagaatactagaagatctagaac    | 1440 |
| Dh | 1381 | AGGAAGAAGAAAAAGATATTTGTGAATGTGTGCTTGCAAAATATATACATCTATGAAAC      | 1440 |

OY 1441 tcaactctaaagtgagcaacacagaacacttgcacagagtaacacagatcatcaggttctga 1500  
 |||||||  
 Db 1441 TCACCTTTAAAGTAGACACACAGAAACTTTGCACAGACTTAACAGTATCAAGTGTGTTAA 1500  
 OY 1501 tgatattatcatagtagttagtcttgactcttggaatatagaagaacacacacttaaaa 1560  
 |||||||  
 Db 1501 TCATATTTGATCTAAGTTAGTTTGTGACTTTGTGGAATGTGAAGAGACACACTTAAAA 1560  
 OY 1561 gaaaagaataaataacagtggttgatccctctcctctgtcttcgcaagtgtccctga 1620  
 |||||||  
 Db 1561 GAAAAAATTAATACAGTGTGGATCCCTTCTCTCTGTTCGCAAGTCTCTGAAAA 1620  
 OY 1621 gactgaacaaaaggaagaatggaattgcacatatcttcgaaagatctgcagaaga 1680  
 |||||||  
 Db 1621 GACTGACAAAAGCAAAAGTGGAAATTCACACATTTCTCGAAGAAAGATTGCCAGGAA 1680  
 OY 1681 tgatacaacagtgagagagcagaattctcctgtataaagagaccaggaactgaaaaa 1740  
 |||||||  
 Db 1681 TCATACACAGTGAAGAGAGCAGAAATTTCTGTATAAGAGACCCAGGAAACTGAAAA 1740  
 OY 1741 gctcctgttatttcaagagccatcccccacccctcaaatgaattgagaggtctaa 1800  
 |||||||  
 Db 1741 GCTCTGTTTATTTAGAGCCCATCCCATCCCTTCAAAATGAATGAGAGGCTTAA 1800  
 OY 1801 gaaaatgagtaataaagtctcaltgttaagtaagctgagatgacatagaagaatt 1860  
 |||||||  
 Db 1801 GAAAAATGAGTAAATTAATGTTCCATGTAACTGACGTGAAGATGACATTAAGACAGATTT 1860  
 OY 1861 taacacagctacactacataaacaacaggaatgacattcttgacattccgaacac 1920  
 |||||||  
 Db 1861 TACACAGCTACTCTACATTAACAAAAACAGGATGCTTTGACATTTCCGAAACAC 1920  
 OY 1921 attaaagtgaatactgaactgaagaactaaggtagatcacataaagtgaacatacag 1980  
 |||||||  
 Db 1921 ATTAAAGTGAATGACTTGAAGAACTAAGGTAGATCACTAATTAATGATACATACAG 1980  
 OY 1981 atctgtacatgtctcgtatctgaactgaacagataatagtgatctcaacccaagaag 2040  
 |||||||  
 Db 1981 ACTGTACATGTTTCTGATTTGATTAACATATATGATGTCTCAACCAACAGAAAG 2040  
 OY 2041 aatactggtccttcttcacgaagaagatcgaagaagaagccttcattcaattcttc 2100  
 |||||||  
 Db 2041 AATATCTGCTGCTTTTCCAGCAAGATCTCAAGCAAAAGACTTTCATTCATTAATTTAC 2100  
 OY 2101 tcatgattctggtctgatacaataaacaagctcacaagaagcacactaactgctcaag 2160  
 |||||||  
 Db 2101 TCATGATTTCTGCTGATTAACATTAACAGTTTCAACAGAGCAACCTTAAGTGTTCAG 2160  
 OY 2161 ggtcctcatcctactcctccctcgaggaaccccaatgaattgacttaagaataatg 2220  
 |||||||  
 Db 2161 GGTCTCATTCCTACTCTCTGAGCAACCCCAATGAATGTGACTTCAAGAAATATGATG 2220  
 OY 2221 ttacactctgttaaaataatcatcgaagaatgaaataatattatagagcgaataa 2280  
 |||||||  
 Db 2221 TTTACTCTTGTGTAATAATCATTCGAAAAAGTGAATAATATTATAGCGAATAATGA 2280  
 OY 2281 aaactctggaacaaatgctgaatttgataaaagaacttgatattatatacaagaaga 2340  
 |||||||  
 Db 2281 AAATCTGGAACAAATGCTGAATTTGATTAAGAAAGAACTGAATTTATACAAAGAA 2340  
 OY 2341 cagaattctgtagtcaacgggtacagctcttaactagactgttccagactagtgaga 2400  
 |||||||  
 Db 2341 CAGAAATTTGATTTACCGGTACAGCTTCTTACTAGACTGTTTCAAGACTAGTGMA 2400  
 OY 2401 atcagaattttgggtttcacagaagctacacagaaagaagtgatataatgcaatttga 2460  
 |||||||  
 Db 2401 ATCAGAAATTTGGGTTTACAACTACACAGAAAGAGTGTATATGCAATTTTATGA 2460  
 OY 2461 taltttggaagaagaatctcagataatctgttaacaagcttcttcctgcccctcaac 2520  
 |||||||  
 Db 2461 TATTGGAAGAGGAATTCAGATTAATCTGTTAAACAGGCTTTTCTCGTCCCTTCAAC 2520

OY 2521 ttacattactggtctttagaatttaaaaaatgacatacttccagaagtataagat 2580  
 |||||||  
 Db 2521 TTTCTACTTTACTGGCTTTTAGAATTTAAAAATGCATCTTTTCCAGAGTATAGAT 2580  
 OY 2581 catatcttgaatttttataaatatgataatgaatcttagatcttaccagctt 2640  
 |||||||  
 Db 2581 CATATCTTGAAATTTTATTAATATGATGAAATTTTACGATTTTTCACACTTT 2640  
 OY 2641 gtttaagaccccaatgataatataatataatcttgcattcttcacagaattga 2700  
 |||||||  
 Db 2641 GTTTACAGACCCCAATGTAATTAATTAATTAATTTTTCGCAATTTTTCACGAAATG 2700  
 OY 2701 taactgttaagaagaataatcagaataaacttgagctgctctgttccattaaaaa 2760  
 |||||||  
 Db 2701 TACCTTTAAAGAAAAATTTACGAAATTAACCTTGTGACTGCTGTTTACATTA 2760  
 OY 2761 aaaaaaaaaaactcgag 2780  
 |||||||  
 Db 2761 AAAAAAAAAAAACTCGAG 2780

RESULT 2  
 AB028070  
 LOCUS AB028070 2719 bp mRNA PRI 04-JUN-1999  
 DEFINITION Homo sapiens mRNA for activator of S phase Kinase, complete cds.  
 ACCESSION AB028070  
 VERSION AB028070.1 GI:4996097  
 KEYWORDS activator of S phase Kinase.  
 SOURCE Homo sapiens Hela CDNA to mRNA, clone:H37.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euteria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS Kumagai,H., Sato,N., Yamada,M., Mahony,D., Seghezzi,W., Lees,E., Arai,K. and Masai,H.  
 TITLE A Novel Growth- and Cell Cycle-Regulated Protein, ASK, Activates Human Cdc7-Related Kinase and Is Essential for G1/S transition in Mammalian Cells  
 JOURNAL Mol. Cell. Biol. (1999) In press  
 REFERENCE 2 (bases 1 to 2719)  
 AUTHORS Masai,H. and Kumagai,H.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-MAY-1999) to the DDBJ/EMBL/Genbank databases. Hisao Masai, Institute of Medical Science, University of Tokyo, Department of Developmental Biology, 4-6-1 Shitokaneda, Minato-ku, Tokyo 108-8639, Japan (E-mail:hisao@ims.u-tokyo.ac.jp, Tel:81-3-5449-5661, Fax:81-3-5449-5424)  
 FEATURES  
 location/Qualifiers  
 1..2719  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /cell\_type="Hela"  
 /chromosome="7"  
 /clone="H37"  
 /map="7q21"  
 518..1222  
 /gene="ASK/ H37 (hubb4)"  
 518..1222  
 /gene="ASK/ H37 (hubb4)"  
 /note="C-terminally truncated due to alternative splicing; an alternative splicing variant of ASK, a regulatory subunit for hucdc7 kinase. This protein could bind to hucdc7 but may not be able to activate its kinase activity."  
 /codon\_start=1  
 /product="activator of S phase Kinase"  
 /protein\_id="BA078327.1"  
 /db\_xref="GI:4996098"  
 /translation="MNSGMRTHSGHFGGCTGQVNNKRPBLSKLTIDNRPESCK PLMGVVFYLDLPSTVISEKLDIDLGQVEEFLSKDLSYLSNKKKEKFAQTIGRI SPVSPESAYTAETSPHPSHDGSSFPDVCLSRGKILVERKAIKDHDFIPSNLSIL NALSWGKTLIHDDIRYVTEQKKRELYLLKSSSTSVRDGKRVGGAQCTFRGRKKP FVKVEMDSGSPAVHLM"  
 2563..2569

polyA\_site



[illegible]

| FEATURES                   | Location/Qualifiers   |
|----------------------------|---|
| Source                     | 1. 2474<br>/organism="Homo sapiens"<br>/db_xref="taxon:9606"  |
| CDS                        | 217. . 2241<br>/note="HscdC7 protein kinase regulatory subunit"   |
|                            | /codon_start=1<br>/product="Dbp4 protein"<br>/protein_id="A041911.1"<br>/db_xref="GI:5306125"   |
|                            | /translation="MNSGAMRHSKGHPGGIOVKKKKRNPRLSKLTNDNRRP<br>PLNGKRVYDLPSVYIASEKIQRIKDLGGVEEPLSKDLSLNKKKEKPFQ<br>SPVPSPSATYAFETSPHPSHDSFSRSPDYVLSKGLLVEKALNDHPIFS<br>NALSWGKYLIIHDIIRYILEQKKRELYLLKRSSTVSDGKSRGSAQKTRIG<br>FVYEDMSOLYRPFYDLITNMPF INVSIKPCSPFVVDKPSRKQATQVRLRG<br>KYGSTIOLKKEKKKKKVECECLOKYEDELDELTLSEQRNFQSNQOVVDD<br>VPPFVEYKNDPKKKKKIKYKSVGLSPVSAVYLKTKQRKRVLELQHSOKDCEQ<br>KEONFLKETQETLEKILPISPIPHPSNELRIKFNKSNKCSMLSTADDIR<br>LPLHKNQECIIIDISEHTLSENDLELRVDHKCNIQASVHVSDETDNSGSG<br>DYVLPFKDLKEDLHSLFPHSDGLYTIINSQSHLYQAKAPHPPEEDNSQ<br>DSLPSGKIHKKVKIILGRNKENLEPNAEDKTEFTIOEENKICSPPOSILL<br>SEKSEFLGFTSTYSEKSGICNVLDIMEVEDSNLLAFPSPSPTSTPTGF" |
| BASE COUNT                 | 888 a 449 c 480 g 657 t   |
| ORIGIN                     |   |
| Query Match                | 88.4%: Score 2458.8; DB 9; Length 2474;   |
| Best Local Similarity      | 99.9%: Pred. No. 0;   |
| Matches 2471: Conservative | 0; Mismatches 2; Indels 1; Gaps   |
| Db 301                     | ctcgtctgtctgttcgagcttcctccctccgagccttggagccggatccggcccggaac 360   |
| Db 1                       | CTCGTTGTGCTTGCGC-CTTCCTCTCCCGCGCTTGAGCCGATCCGCGCGGAAC 59  |
| Qy 361                     | ccgacctgcagacgcgcgtacctactcgtcgtatgagcccgtagctcgcgcgaagaagaagg 420  |
| Db 60                      | CCGACCTCCACACGCGGACCTCTACCTCCAGAGCGCGTACGCTGCGGCGAAGGAGAGG 119  |
| Qy 421                     | cggccgcctccttcacaacgacgcggggggagccgctccttcgcgcgtgcgcgcgtgcgaact 480   |
| Db 120                     | CGGCGCTCTCTGCACACAGCGCGGGGGGAAGCGCTCTTCCGCGCTGCCCGGTGCGACACT 179  |
| Qy 481                     | ttctccggaaccacagatgtgaigtgcgcggcgacgtccatgaactccggagcccatgaagat 540   |
| Db 180                     | TTTCTCGGAGCCAGAGATGTAGTGTCCGGGCGACATGCCATGTACATCCGAGCATGAGAT 239  |
| Qy 541                     | ccacagttaaggaacttccaggggtgaaatccaaatcaaaaaagaaaaaacagacatc 600  |
| Db 240                     | CCACAGTAAAGACATTTTCCAGGGTGGATCCAAAGTAAAAATCAAAAAACACACATC 299   |
| Qy 601                     | tctgaatcctctgaaanaactgatacagacgcgaagaaatccaatgtaagccaacttgggg 660   |
| Db 300                     | TCTGAATCTCTGAAAATCTGTAACAGGCCGAAAAATCCAAATGTAAGCCACTTTGGGG 358  |
| Qy 661                     | aaaagtatttacccttgacttacccttcgtaaccatalcgtgaaanaacttcaaaaagacat 720  |
| Db 360                     | AAAAGTATTTTAACTGACTTACCTTCCTGTGCACATATCTGAAAAAATCTCAAAAGGACAT 419   |
| Qy 721                     | taaggaatcctggagggcgagtttgaagaatcttctaagcaagaatacaagttatcttatttc 780   |
| Db 420                     | TAAAGATCTGGAGGCGCAGTGAAGAAATTTCTCAGCAAAAGATATCTAGTTATCTTATTTTC 479  |
| Qy 781                     | aaataagaagaagaacttaaatgtgacaaaacttgggtcgaaattctctctgtaccaaatcc 840  |
| Db 480                     | AAATAGAGAGAGCTAAATTTTGACAAACCTTGGGTGCAATTTCTCTGTACCAAGTCC 539   |
| Qy 841                     | agaatctgatalactgcgaagaacacacttcaatccatcccaagccatgaltgaagattcat 900  |
| Db 540                     | AGAACTCTCATATTACTGCAAGAACCATCTCACTCATCCAGCCATGATGGAAGTTCATT 599   |
| Qy 901                     | taagtaccagaacacagtgtctttaagcagaagaataattatagtgtgaaaaagctatcaa 960   |
| Db 600                     | TAAAGTACCACACACAGTGTCTTTAAGCAGGAAATTTATTAAGTAAAAAGCTATCAA 659   |

|    |      |  |      |
|----|------|--|------|
| QY | 961  | ggacacatgatttattccctcaaatgatatattatcaaatgcctgtccatggggaataa          | 1020 |
| QY | 962  | ggacacatgatttattccctcaaatgatatattatcaaatgcctgtccatggggaataa          | 1020 |
| Db | 660  | GGACCATGATTTTATTCCTCCAAATAGATTTATTCAAATGGCTTGCATGGGAGTAAA            | 719  |
| QY | 1021 | aattcctcatatgtagacataagaatactacatactgaaacaaagaaagagttgattt           | 1080 |
| Db | 720  | AATTCATATTTGATGACATATTTGATCTACATATTTGACAAAGAAAAGATTGTTT              | 779  |
| QY | 1081 | actcaagaataatcaagtaacttactcagtaagaagatcggggcaaaagattgttagtcaca       | 1140 |
| Db | 780  | ACTCAAGAAATCAAGTACTTTCGTAAGAGATGGGGGCAAAAGATTGGTACTGGTGAC            | 839  |
| QY | 1141 | aaaaacaagaacaggaagacatcaaaaagccctttgttaagttggaagatatgagcaact         | 1200 |
| Db | 840  | AAAAACAAGAACAGAGAAAGCTAAAAAGCCCTTTGTAAAGGTGAAAGATATGACCCTAACT        | 899  |
| QY | 1201 | ttatagagccattatcatcttcagctgacccaataatgaccttataaatcttcatcttcagaa      | 1260 |
| Db | 900  | TTATAGGCCATTTTATCTCAGCTGACGACCAATATGCCCTTTATTAATTTATCTATTCGAA        | 959  |
| QY | 1261 | ggccctgagctccatttgaatgtagaacagcaatcagatgtcgaacaaagaaactcagtttaa      | 1320 |
| Db | 960  | GGCCTGAGTCCATTTGATGTGACAGCCATCTAGATNTCCAAAGCAAACTCAGGTAAA            | 1019 |
| QY | 1321 | acttaagaatccaacagatgagcagaataagtatggtgagacccatcatccaactcagttgaa      | 1380 |
| Db | 1020 | ACTAAGATCCAAACAGATGGGATATGATATGTGGAACTCATATTCAACTCCAGTTGAA           | 1079 |
| QY | 1381 | aaagaagaagaaaaaagatgatatgtgaaatgttgccttcagaataatgaaagatccagaac       | 1440 |
| Db | 1080 | AGAGAAAGAAAAAAGATTTGTGAATTTGCTGCAGAAATATGAAGATCTGAAAC                | 1139 |
| QY | 1441 | tcaactcttaagtgagaaacagagaactctgcacagaatgaaccagtatcagaattgtga         | 1500 |
| Db | 1140 | TCACCTCTCAAGTATGACACACAGAAACTTTTGCCACAGATGATACCAATATCAAGTTGTGA       | 1199 |
| QY | 1501 | tgatatgtatcctcaagtttagcttttgactctgttggaatatgaaagagaacaccttaaaa       | 1560 |
| Db | 1200 | TGATATTTGATCTAAGTTGATGATTTTCTACTTTGTGGATATGTAAAGAGACACCTTAAAA        | 1259 |
| QY | 1561 | gaaagaataaataatcagctgttggaatcccttctctcgttcttcgcaagtgtccgtgaana       | 1620 |
| Db | 1260 | GAAAGAAATAAAAATACAGTGTGGATCCCTTTCCTCTGTTCGCAAGTCTCTEAAAAA            | 1319 |
| QY | 1621 | gactcgacaaaggaanaaagtggaaatctgcaacataatctcccaagaagatctgcagaaga       | 1680 |
| Db | 1320 | GACTGACAAAGGAAAGATGGAATTTGCAACATTTTCTCAGAAAAGATTGCCAGGAAGA           | 1379 |
| QY | 1681 | tgtatacaagtgaaagagcagaattctctgtataaagagaccagaagaaactgaaaaaaa         | 1740 |
| Db | 1380 | TGATACACAGTGAAGGAGCGAGATTTCTGTATTAAGAGACCCAGAGAACTGAAAAAAA           | 1439 |
| QY | 1741 | gctccggttatttcaagagcccatcccccacaccttcaatgaatatgaaagggcttaagaa        | 1800 |
| Db | 1440 | GCCTCGTTTATTTCAAGAGCCATCCCGCACCTTCAATGAATTTGAGGGCTTAAAGA             | 1499 |
| QY | 1801 | gaaatagatataaatggttccatgfttaagatacagctgaaagatgaaatgaaacagaattc       | 1860 |
| Db | 1500 | GAATATGATTAATTAATTTCCATGTTAATGATACCTGAACATGATACATTAAGACAGATTT        | 1559 |
| QY | 1861 | tacacagctacctctacataaaaaacaaacagaatggaatttgaacatttcgaaacacac         | 1920 |
| Db | 1560 | TTACACAGCTACTCTTACATTAATAAACACAGAAAGCACTTTTGACATTTCCGAAACAC          | 1619 |
| QY | 1921 | attaaatgaaataatgacttaagaagaactaagggttagataactataatgtcaacatacagc      | 1980 |
| Db | 1620 | ATTAAATGAAATATGACTTGAAGAAACCTAAGAGGTGAATACATCATTAATATGTACATACGC      | 1679 |
| QY | 1981 | atctgtacatgcttctgtcttccatgacagataatgtagatctcaaccacaaacagaatgc        | 2040 |
| Db | 1680 | ATCTGTACATGTTTCTGTCAATTTCACTATACATAGATATATAGTGAATCTTAACCAAAACAGAAATG | 1739 |

|    |      |   |      |
|----|------|---|------|
| QY | 2041 | agctactgtgcttcttcacgcaaaagatcttaagaaaaaggaactcatttcaatcttc      | 2105 |
| Db | 1740 | AGATACCTGGCTTTTCCAGAAAGATCTCAGAGAAAGACCTTCATTCAATTTTAC          | 1799 |
| QY | 2101 | tcatgtatctgtctgtataacaabaacagcttcacaagagcaactaactggttcagcaaa    | 2160 |
| Db | 1800 | TCATGTATTCGTGTGATTAACAATAAATCAAGTTCCACAGACCTTAACCTTCACGCAAA     | 1859 |
| QY | 2161 | ggtctcatctccatctctctctctgtagagaccatgtaatgtaacttcaagataatggatag  | 2220 |
| Db | 1860 | GGCTCCATTCCTACTCTCTCTGAGAACCCATGAATGTACATTCAAGATATGTGATAG       | 1919 |
| QY | 2221 | tttaacctctctgtataaatcacatcgaagaagtgagaaataatataagagaataatagaaga | 2280 |
| Db | 1920 | TTTATCCCTCTGTATAAATTCATCCAAAAGTAAATATATATTAGACGAATGTGAATAAA     | 1979 |
| QY | 2281 | aaatctggaaccaaatgctgtgaattgtataaagaactgtaatttacaacaagaagaa      | 2340 |
| Db | 1960 | AAATCTGGAACCAATGCTGTAATTTGATTAATAAACACTGAATTTATACACAGAAAGAAA    | 2038 |
| QY | 2341 | caggaattttagtctaacccggaagactcttactagaactggttcaagactgataagaga    | 2400 |
| Db | 2040 | CAGATTTGTAGTTCACCGGTRACACTTTACTAGACTTGTTCAGACTGTGAAGAGAA        | 2099 |
| QY | 2401 | atcagaatttltgggttttcaacaagctctcaagaagaagtgatataagaattttta       | 2460 |
| Db | 2100 | ATCCGAATTTTGGGTTTCACACACTCACAGAAAGAGTGTATATGCCAATGTTTAA         | 2155 |
| QY | 2461 | tatttgggaagagaaatctcaagaatactgtttaacagcgttttctcgtccctctcaac     | 2520 |
| Db | 2160 | TATTTGGGAAGAGCAAAATCAGATTAATCTGTTAAACGGCTTTCTGTCCTCCCTCAAC      | 2215 |
| QY | 2521 | ttcacaatttaacgagcttcttgaatttcaaaaaatgataactttccaagaagtataagat   | 2580 |
| Db | 2220 | TTTCACATTTTACTGCGCTTTTGAATTTAAAAAATGCATCTTTTCAGAAAGATAAAGAT     | 2275 |
| QY | 2581 | catctcttgaaaatttttaataataatfatcgaagaatctctgaagatttttttaccagctt  | 2640 |
| Db | 2280 | CATATCTCTAAATTTTATTAATPTGATAGGAAATTTCTAGATTTTTCACAGCTTT         | 2335 |
| QY | 2641 | gttttagagcccaatgataaatataaataaataatcttggcaatttctacaagaattgaa    | 2700 |
| Db | 2340 | GTTTACAGACCCAAATGTAAATTTAAAAAATAATTTTGCATTTTCTACAGATGTAA        | 2395 |
| QY | 2701 | tactgtttaagaaaaaatacaagaataactgtgacgtgtctgttttataataaaaaa       | 2760 |
| Db | 2400 | TACCTGTTAAAGAAATTTTACAGATTAACCTTGACGTGTCTGTTTTACATTAATAAAA      | 2455 |
| QY | 2761 | aaaaaaaaaaaaa 2774  |      |
| Db | 2460 | AAAAAAAAAAAAA 2473  |      |

|            |  |            |                     |             |
|------------|--|------------|---------------------|-------------|
| RESULT     | 4  |            |                     |             |
| AF160876   |  |            |                     |             |
| LOCUS      | AF160876   | 2460 bp    | mRNA                | PRI         |
| DEFINITION | Homo sapiens DBP4-like protein (DBP4)  |            | mRNA, complete cds. | 22-JUL-1999 |
| ACCESSION  | AF160876   |            |                     |             |
| VERSION    | AF160876.1   | GI:5566280 |                     |             |
| KEYWORDS   | .  |            |                     |             |
| SOURCE     | human.   |            |                     |             |
| ORGANISM   | Homo sapiens   |            |                     |             |
|            | Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;   |            |                     |             |
|            | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  |            |                     |             |
| REFERENCE  | 1 (bases 1 to 2460)  |            |                     |             |
| AUTHORS    | Hollingsworth,R.   |            |                     |             |
| TITLE      | Use of a semi-automated yeast two-hybrid system to identify proteins that interact with the human Cdc7 protein |            |                     |             |
| JOURNAL    | Unpublished  |            |                     |             |
| REFERENCE  | 2 (bases 1 to 2460)  |            |                     |             |
| AUTHORS    | Hollingsworth,R.   |            |                     |             |
| TITLE      | Direct Submission  |            |                     |             |
| JOURNAL    | Submitted (19-JUN-1999) Genomic Sciences, GlaxoWellcome, Inc., Five  |            |                     |             |

Moore Drive, Research Triangle Park, NC 27612, USA

FEATURES  
Source Location/Qualifiers

1..2460  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="7"

/map="7q21"

1..2460

/gene="DBF4"

200..2224

/gene="DBF4"

/codon\_start=-1

/product="DBF4-like protein"

/protein\_id="AAD4537.1"

/db\_xref="GI:556281"

/translation="MNSGAMRIHSKHFGOGIIVKRNKRLPSLKLTDNRPEKSKC

PLMGKVFYLDLPSVTSISEKLQIDKIDGLRVEEFLSKDYSITSKKEKFKQTLGRI

SPVSPESAYTAETTSPPSHDSSEFSPVCLSRGLKLVKAIKDHDFIPSNILS

NALSMGVKLIHIDDI RYITEOKKKELYLKSTSPVROGKRGAKOQTPRGRLKP

FKVEDMSQLRREPLYLTNNPFTYSTQKPSPPDVDRPSMOKOTYKRLIOTDGD

KYGTSTIOLQLEKKKGYCECLOKTYEDLEHLISEDRNFAGSNQYOVDDIVSKL

VEFVEYERKDPPEKKRIRYSGVSLPVSAVLKTEOKKEVLAHI SQDCQEDDTV

KEONFLYKETQTEKKLFISEPIHPNSNELGLNKKSNKCSMLSTADDIROMFTQ

LPLHKKQECILIDISEHTLSEMDLELRVDYKCNIOASVHSDSTDNSSGSPORS

DVLYEPARKDEKIDHSLIETHDSGLITINSSEHILVYAKAPPHNPREPNCDPKM

DSLPSGKIHKKYKIIILGRNKRNEPNAEPDKRTFTTQENRITSSPIQSILDLFOT

SEKSEFEGFTSYTEKSGICNVLDIWEENSDNLTAFFSSPSTFTGF"

BASE COUNT 891 a 445 c 475 g 649 t

ORIGIN

Query Match 88.3%; Score 2456; DB 9; Length 2460;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0;

Matches 2456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 cttccctcccgcccttgagccgagatccggcccgaaaccgacccctgcaagcggt 378  
1 cttccctcccgcccttgagccgagatccggcccgaaaccgacccctgcaagcggt 60  
QY 379 accctactcgtagagccgtagctgcggaagagagagcgccgctctgcaag 438  
61 accctactcgtagagccgtagctgcggaagagagagcgccgctctgcaag 120  
QY 439 gccgggggaagccggtcttcgaggtgcccgggtgacacatttcccgaccagcatg 458  
121 gccgggggaagccggtcttcgaggtgcccgggtgacacatttcccgaccagcatg 180  
QY 499 taagtccgggagcagctgacatcgacccggaagcagatccacagtaagacatt 558  
181 taagtccgggagcagctgacatcgacccggaagcagatccacagtaagacatt 240  
QY 559 ccagggtggaatccaaatgcaaaaaacagacacatctctgaaatctctgaaac 618  
241 ccagggtggaatccaaatgcaaaaaacagacacatctctgaaatctctgaaac 300  
QY 619 tgaatacagagccagaaaaatccaaatgtaagcactttggggaaaagtatttacttga 678  
301 tgaatacagagccagaaaaatccaaatgtaagcactttggggaaaagtatttacttga 360  
QY 679 cttacctctgtcaacatctgaaaaacttcaaaaggaatcaagatctgggaaggcg 738  
361 cttacctctgtcaacatctgaaaaacttcaaaaggaatcaagatctgggaaggcg 420  
QY 739 agttgaagaattctcagaagaataatcagttatcttatttcaataaagaagaactaa 758  
421 agttgaagaattctcagaagaataatcagttatcttatttcaataaagaagaactaa 480  
QY 799 atttgacaacacttggtgcgaattctcctgfracaagccgaatctgcataactgc 858  
481 atttgacaacacttggtgcgaattctcctgfracaagccgaatctgcataactgc 540  
QY 859 agaaacacacttcaaccatccagcagatggaagtcttaagttcaaccagacagat 918  
|||||

DB 541 AGAAGCAGCTTACCTCATCCAGCCATGATGGAAGTTCATTTAAGTCACACAGACAGT 600  
QY 919 ggtttaagcagaggaataatattgattgaaaaagctatcaagacatgattatctcc 978  
DB 601 ggtttaagcagaggaataatattgattgaaaaagctatcaagacatgattatctcc 660  
QY 979 ttcaaatagatatttcaaatcccttgatcagggaggaataaatcttcaattatgta 1038  
DB 661 ttcaaatagatatttcaaatcccttgatcagggaggaataaatcttcaattatgta 720  
QY 1039 catatagatacatatgaaacaaagaagaagaagtgatatttactcaagaatcaagttac 1098  
DB 721 catatagatacatatgaaacaaagaagaagaagtgatatttactcaagaatcaagttac 780  
QY 1099 ttcaagtaagaatgaggggcaaaagattggttagtgglygcaaaaaaacaagaacaggaag 1158  
DB 781 ttcaagtaagaatgaggggcaaaagattggttagtgglygcaaaaaaacaagaacaggaag 840  
QY 1159 actcaaaaagcccttggtaaggtggaagatagagcccaattatagccattatct 1218  
DB 841 actcaaaaagcccttggtaaggtggaagatagagcccaattatagccattatct 900  
QY 1219 tcaagctgaaccaatagccctttaaattatctatctcaagaagccctgagctcattga 1278  
DB 901 tcaagctgaaccaatagccctttaaattatctatctcaagaagccctgagctcattga 960  
QY 1279 tgaagaacagccatctagatctgcaaaagcaactcaggtttaaactaagaatccaacaga 1338  
DB 961 tgaagaacagccatctagatctgcaaaagcaactcaggtttaaactaagaatccaacaga 1020  
QY 1339 tggcgaatagatgagtggaagcccaattcaactcaggttgaagaaggaagaagaagaag 1398  
DB 1021 tggcgaatagatgagtggaagcccaattcaactcaggttgaagaaggaagaagaagaag 1080  
QY 1399 atatttgaattgtctctgaggaataatgaaatctcaagaactcacccttcaagtga 1458  
DB 1081 atatttgaattgtctctgaggaataatgaaatctcaagaactcacccttcaagtga 1140  
QY 1459 acaagaaactttgcaagagatgaacagatccaagtgtgtagatattgtatctaagt 1518  
DB 1141 acaagaaactttgcaagagatgaacagatccaagtgtgtagatattgtatctaagt 1200  
QY 1519 agttttgactttgtggaataatgaaaggaacacacccaataaagaagaataaataacag 1578  
DB 1201 agttttgactttgtggaataatgaaaggaacacacccaataaagaagaataaataacag 1260  
QY 1579 tgttgaatcccttctcctgttctgcaagtgctcctgtaaaagactgaaacaaagaagaa 1638  
DB 1261 tgttgaatcccttctcctgttctgcaagtgctcctgtaaaagactgaaacaaagaagaa 1320  
QY 1639 agtggaaattgcaacataatctcagaagaattgccaagaagatgatacaacagtgaaaga 1698  
DB 1321 agtggaaattgcaacataatctcagaagaattgccaagaagatgatacaacagtgaaaga 1380  
QY 1699 gcagaattctctgataaagaagccaggaactgaaaaaaactcctgttatttcaaga 1758  
DB 1381 gcagaattctctgataaagaagccaggaactgaaaaaaactcctgttatttcaaga 1440  
QY 1759 gcccatcccccaccccttcaaatgaaatlgagaaggcttaatlgagaataatgagtaaatg 1818  
DB 1441 gcccatcccccaccccttcaaatgaaatlgagaaggcttaatlgagaataatgagtaaatg 1500  
QY 1819 ttccatgttaagtaagcctgaaatgacataaagaacgaatttcaacagctacccttaca 1878  
DB 1501 ttccatgttaagtaagcctgaaatgacataaagaacgaatttcaacagctacccttaca 1560  
QY 1879 taaaaacaaacaggaatgagcttcttgaaacatttcgaaacacacattgaatgaatgaact 1938  
DB 1561 taaaaacaaacaggaatgagcttcttgaaacatttcgaaacacacattgaatgaatgaact 1620  
QY 1939 agaaagaactaaggtagataactataaatgtaacatacaggaactctgtaatttctga 1998  
DB 1621 agaaagaactaaggtagataactataaatgtaacatacaggaactctgtaatttctga 1680  
|||||

|             |   |   |      |
|-------------|---|---|------|
| Qy          | 1999  | tttcagtcacagtaataagtcgactctcaacccaacaaagtcgaatacctcgtctttcc       | 2058 |
| Db          | 1681  | tttccattacagatattatgtgatctctcaacaaaacagaaatgcatactgtctttcc        | 1740 |
| Qy          | 2059  | agcaagaagatctcaagaagaagactctattcaatcttcacatgattctggtctat          | 2118 |
| Db          | 1741  | agcaaaagattctcaagaaaagaaagactcttattcattattttactatgatctggtctat     | 1800 |
| Qy          | 2119  | aacataaaccagttccacaagaagcaactaacctgttcaggcaaaagctccattccatcc      | 2178 |
| Db          | 1801  | aaacattaaacagttccacaagaagcaactaaccttattcattattttactatgatctggtctat | 1860 |
| Qy          | 2179  | tcctgaagaaccacaatgaatgtgacttccaagataatgatatgattactctctgtitaaat    | 2238 |
| Db          | 1861  | tcctgaggaaccaccaatgaatgtgacttccaagataatgatatgattactctctgtitaaat   | 1930 |
| Qy          | 2239  | acatcgaaaagtcgaataataatctagaacgaataatgaaaagaataatcttgaaacccaatgc  | 2298 |
| Db          | 1921  | acatcgaaaagtcgaataataatctagaacgaataatgaaaagaataatcttgaaacccaatgc  | 1980 |
| Qy          | 2299  | tgaattctgataaagaacatgaattcttccacaagaagaagaacgaattctgattctacc      | 2358 |
| Db          | 1981  | tgaattttgattaaaagaaactgaattttttacacaaagaaacgaattttgattctacc       | 2040 |
| Qy          | 2359  | ggtacagccttactgaactgtttctagaactgtagaagaagaatccgaatttttgattt       | 2418 |
| Db          | 2041  | ggtacacgcttttactgaactgttttgaactgtttgaactgtttgaactgttttgattt       | 2100 |
| Qy          | 2419  | cacaagctcacacagaagaagaagatggtatatgcaatgcttttagatatctgggaagaggaaa  | 2478 |
| Db          | 2101  | caaaagctcacacagaagaagaagatggtatatgcaatgcttttagatatctgggaagaggaaa  | 2160 |
| Qy          | 2479  | cttcagataatcgtttaaacagcgcttttctcgtcccttcaactcttcaatttacttgctt     | 2538 |
| Db          | 2161  | tttcagatattatcgtttaaacagcgcttttctcgtcccttcaactcttcaatttacttgctt   | 2220 |
| Qy          | 2539  | ctagaattctaaanaatgcgaactcttcagaagtcgaatgcatatctcttgaatctt         | 2598 |
| Db          | 2221  | tttagaatttttaaaaaatgcgaactcttcagaagtcgaatgcatatctcttgaatctt       | 2280 |
| Qy          | 2599  | ataaataatgatctgaaatctctttagaattcttcttaccagccttgtttacagacccaatgct  | 2658 |
| Db          | 2281  | atfaaataatgatctgaaatctctttagaattcttcttaccagccttgtttacagacccaatgct | 2340 |
| Qy          | 2659  | aaattctaaanaatgaattctgcaattctctcagaatcttgataccgtttaagaagaat       | 2718 |
| Db          | 2341  | aaatttttaaaaaatgaatttttgcgaattttcttcagaatcttgataccgttttaagaagaat  | 2400 |
| Qy          | 2719  | ttcagaataaactgttgactggtcgtgttttcataataaaaaaaaaaaaaaaaaa           | 2774 |
| Db          | 2401  | ttcacaaataaactgttgactggtcgtgttttcataataaaaaaaaaaaaaaaaaa          | 2456 |
| RESULT      | 5   |   |      |
| ACOL18977/c |   |   |      |
| LOCUS       | ACOL18977   | 236236 bp   | DNA  |
| DEFINITION  | Homo sapiens chromosome 10 clone RP11-386C23, WORKING DRAFT       |   |      |
| SEQUENCE    | SEQUENCE, 12 unordered pieces.                                    |   |      |
| ACCESSION   | ACOL18977   |   |      |
| VERSION     | ACOL18977.6   |   |      |
| KEYWORDS    | HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFIN.       |   |      |
| SOURCE      | human.  |   |      |
| ORGANISM    | Homo sapiens  |   |      |
| REFERENCE   | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |   |      |
| AUTHORS     | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.        |   |      |
| TITLE       | Smith,D.R.  |   |      |
| JOURNAL     | Genome Therapeutics Corporation Sequencing Center: Human Genome   |   |      |
| REFERENCE   | Unpublished   |   |      |
| AUTHORS     | 2 (bases 1 to 236236)   |   |      |
|             | Smith,D.R.  |   |      |

```

TITLE      Direct Submission
JOURNAL    Submitted (25-Dec-1999) Genome Therapeutics Corporation, 100 Beaver
           Street, Waltham, MA 02453, USA
COMMENT    On Jul 26, 2001 this sequence version replaced g1:14091812.

-----
Genome Center
Center: Genome Therapeutics Corporation
Center code: GTC
Web site: http://www.genomecorp.com/
Contact: gtc-seqcenter@genomecorp.com
-----
Project Information
Center project name: hg092
-----
Summary Statistics
Sequencing vector: N/A
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap: version 990315
Consensus quality: 228735 bases at least Q40
Consensus quality: 230209 bases at least Q30
Consensus quality: 231169 bases at least Q20
Insert size: 235235; sum-of-contigs
Quality coverage: 6.3x in Q20 bases; sum-of-contigs
-----

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1      1146: contig of 1146 bp in length
*
*      1147      1246: gap of unknown length
*
*      1247      2466: contig of 1220 bp in length
*
*      2467      2566: gap of unknown length
*
*      2567      4025: contig of 1459 bp in length
*
*      4026      4125: gap of unknown length
*
*      4126      5165: contig of 1040 bp in length
*
*      5166      5265: gap of unknown length
*
*      5266      6309: contig of 1044 bp in length
*
*      6310      7552: contig of 1143 bp in length
*
*      7553      8716: gap of unknown length
*
*      8717      8816: contig of 1064 bp in length
*
*      8817      10307: gap of 1451 bp in length
*
*      10308      10407: gap of unknown length
*
*      10408      11813: contig of 1506 bp in length
*
*      11914      12013: gap of unknown length
*
*      12014      13363: contig of 1350 bp in length
*
*      13364      13463: gap of unknown length
*
*      13464      106707: contig of 93244 bp in length
*
*      106708      106807: gap of unknown length
*
*      106808      236336: contig of 129429 bp in length.

FEATURES
        source
            1..236336
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="10"
                /clone="RP11-386c23"
                /clone_11b="RPC1-11"
            1..1146
                /note="assembly_name:Contig7"
            1247..12466
                /note="assembly_name:Contig11"
            2567..4025
                /note="assembly_name:Contig12"
            4126..5165
                /note="assembly_name:Contig13"
            5266..6309
                /note="assembly_name:Contig16"
            6410..7552
                /note="assembly_name:Contig17"
            7653..8716
                /note="assembly_name:Contig18"

```



[illegible]

|    |        |   |        |
|----|--------|---|--------|
| Db | 217760 | GATGACATTTACATCTCTACATTGAACAAAAGAAAAA - AGTTGCATTTCTACCAAGAAATCA  | 217703 |
| OY | 1094   | agctctcaagcaagaagatctggggcaaaaagcttgatagctgtgcacaaaacaagaaca      | 1153   |
| Db | 217702 | AGTACTTCAGTAAAGATGGGGGCAAAAGAGCTTGGTGTGCTGCACAAAACAAACAAACA       | 217643 |
| OY | 1154   | ggaagactcaaaaagccctttgttaaggctggaagatacgaccacttataagccatt         | 1213   |
| Db | 217642 | GGAAAGACTCAAAAAGCCCTTTGTAAAGGTGGAGATTTGAGCCACCTTTATAGGCATTT       | 217583 |
| OY | 1214   | tatctcaagctgaccatatacgcttttatataattatctcttaagaagcccttcactca       | 1273   |
| Db | 217582 | TATCTTCAGCTGACCCAATATGCCCTTTATTAATTTATTTCAATTGAGAAAGCCCTGCACCTCA  | 217523 |
| OY | 1274   | tttgtatgagacaagccactcagatgatcgaagaagcaaacctaggtttaaactaaagctcaa   | 1333   |
| Db | 217522 | TTTGATGTAGCAAGCCACTCTAGTATGCCAAAAGCAAACTCAGGTTAACTAAATTCCAA       | 217463 |
| OY | 1334   | acagatgycgaataagatagctgctggaacctcaattcactccagtttgaaagaaagaa       | 1393   |
| Db | 217462 | ACAGATGCGCATTAAGTATGGTGGCAACTCAATTCACCTCCAGTTGAAAGAGAGAGAAAA      | 217403 |
| OY | 1394   | aaaggaatctgtgaatctgtcctctgcagaaataatgaagaacctaaagctccctcaagt      | 1453   |
| Db | 217402 | AAAGGATTTTGTGAATGTGCTTCCACAAAATGCAAAATGTAGAAAT-----               | 217353 |
| OY | 1454   | gagcacacagaagaacttgacacagagtaaacagatcaagttgtgtatattgtatct         | 1513   |
| Db | 217354 | -----   | 217353 |
| OY | 1514   | aagtagtttttgactgtggaatatgaaaaaggacacctaaaaaagaagaataaaa           | 1573   |
| Db | 217354 | -----ATAAAAA-----   | 217348 |
| OY | 1574   | tacagtgcttgatcccttctccgtttcttgcagtgctctgaaaaagactgaaacaaaag       | 1633   |
| Db | 217347 | TACAGTGTGTGATCCCTTTCTCCTGTTTCTGCAAGTCTCTGAAAAAGCTGAACGAAAG        | 217288 |
| OY | 1634   | gaanaagctggaatctgcaacatattctctcaagaagaatctgcagagaagaatgaataacagtg | 1693   |
| Db | 217287 | GAAAAAGTGGAAATGCCAATATTTCTCAGAAANATTTGCCGGGAAGATGATATACGGTG       | 217228 |
| OY | 1694   | aagsgcsgaaattctcgtlatalaaagagacccagagaacctgaaaaaagctccgttatt      | 1753   |
| Db | 217227 | ATAGAGCGAATTTCTGTATAAAGAGACCCAGGAAACTGAAAAAA - CTTCTGTTTATT       | 217169 |
| OY | 1754   | tcaagagcccatcccccaccccttcaaatgaaattgagaggtcttaatggaanaatgagtaat   | 1813   |
| Db | 217168 | TCAGAGCCCATCCCCACCCCTTCAAAATGACTTGAGGGCTTTAAACAAGAAATGAGATAAT     | 217109 |
| OY | 1814   | aaatcttcagttaagtaagctcgaagatgaaataagaaatgaagaaatttacaagctaacct    | 1873   |
| Db | 217108 | AAATGTTCCATGTTTAAGTACAGCTGAGAGATATCAATAAGACAGAAATTTTACACAGCTACT   | 217049 |
| OY | 1874   | ctacataaaaacaacaagaaatgaaatctcttgacaatttcgagacacacattaaatgaaat    | 1933   |
| Db | 217048 | CTACATTAATAAACAAACAGAGATGCAATCTTCTGTACATTTCCGACACACATTAAGGAAAT    | 216989 |
| OY | 1934   | gaactagaagaactaaggctagatcactataaatgtaacatcacaggcaatctgtacatgtc    | 1993   |
| Db | 216988 | GACTTACAAAGAACTTAAGATGTAGTAACTTAATTAAGTAACTATACAGCATCTGTTCACTGT   | 216929 |
| OY | 1994   | tctgatttcagtaacagataaataagtgatcttcaaccanaaagaagatcaagtactgtgctt   | 2053   |
| Db | 216928 | TCTGATTTCTGTACAGATAATTAATGTGCAATCTCAACAAAACAAAAGTCCAGATCCTGTGCTT  | 216869 |
| OY | 2054   | ttccagcaaaagatctcgaagaagaagacctcattcaatacttaactgaatctcgtgt        | 2113   |
| Db | 216868 | TTTCCAGCAAAAGATCTCAAGAAAGAGACCTTCATTCATATTTACTCATGATTTCTGGT       | 216809 |
| OY | 2114   | ctgataaacaataaagcttcaagaagcaacctaaagttcgaagcaaggtctcaatccat       | 2173   |



|    |      |   |      |
|----|------|---|------|
| Oy | 1050 | accttgcacaaagaagaaaggttcttattctctcaagaatccaagtacttcagtaag       | 1109 |
| Db | 727  | ACATTTGAACGAAGAAGAAAGAGTATTGTTCACCTCAAGAAACAAGTACTTGTGAAGG      | 786  |
| Oy | 1110 | atggaggcaaaagacttctgttagtggcgccaaazaaacaagacaggaagactccaaaagc   | 1169 |
| Db | 787  | ATTGGGAAAAAAGCTGGCAGCTACAAATACAAAAGGCCAGAACGGAGACTCAAAAAGC      | 846  |
| Oy | 1170 | ctttctgaagagtgagaaatatgagccaacttataagagcatttattctcagctgacca     | 1229 |
| Db | 847  | CTTTTAAAGGTTGAAGATGTGAACCGGCTTTACAGGCACTTCACTTCAGCTGACCA        | 906  |
| Oy | 1230 | atatgccttataaattattctattcagaagccctgcagtcacatttgatgbagaacg       | 1289 |
| Db | 907  | GTGGCTCTTAACTAATCTATGTCACGACACAGCCTTGACGCCATTGACATGAAGAAGC      | 966  |
| Oy | 1290 | cattctagtatgcanaagccaactcaggttcaactaagaatccaacagatggcgtaat      | 1349 |
| Db | 967  | CATCACACGTCACAAAGCAAGCTCACGCTTAACCCAGACTTAATCGGATGGTGAAT        | 1026 |
| Oy | 1350 | atggtgaaacctcaattcaactcaggttgaagaagagaagaaaagatacttgat          | 1409 |
| Db | 1027 | GTGGTGAAACCCCGATTCACTCAATTGAGAGAAAGGAAAAAAGGATATTGTGAT          | 1086 |
| Oy | 1410 | gttccttcagaaataatgaagatctagaactcacttccttaagttagcagacagaaact     | 1469 |
| Db | 1087 | GTTCCTTACAGAAATACGAAGATCTTGAAATCACCTTTGTAGTGAGAAACCAAAACT       | 1146 |
| Oy | 1470 | ttgcacagagtaaccagtatcaagtgttgatgatatgtatctagttagtllttgact       | 1529 |
| Db | 1147 | TGCGACAGAGATACACAGTATCAAGTTGTGATGATATTATCTAAGTTAGTTTGTAT        | 1206 |
| Oy | 1530 | tttggagaatctgaaaagaacacaccccaaaaagaagaaataaacatcacagtgttgatcc   | 1589 |
| Db | 1207 | TTTGGAATTTAGAGAGACACACTTAATAAGAAATAAATATACGCGTGGCTCT            | 1266 |
| Oy | 1590 | tttctccgtgttcctgcagtgctctgtaaaaagactgaacaaaggaanaagtggatgc      | 1649 |
| Db | 1267 | TTTCTCTGTTCTTGCAAAAGTCTCTGAAAAACCTCAACCAAGGAGAAAGCTACAGTTGG     | 1326 |
| Oy | 1650 | aacatllctctcagaagaagatgccaggaagatgatcaaacagtgaaagcaga-----      | 1703 |
| Db | 1337 | AGGCCATTTTCCAGAAAGACATGATGTAAGTAATGACATCGCTGAGAGATTTTTC         | 1386 |
| Oy | 1704 | -----attccgtataaagaagaccccgagaaactgaaaaaagctcctgttattt          | 1755 |
| Db | 1387 | AGTGTGAAGATATCCACTGTGGAAGATATCCAGAAACCTCAACAAAGCGCTTGTTGGCTT    | 1446 |
| Oy | 1755 | cagagcccatcccccaccaccttaaatgaaatctgagagggcttaagaanaatgagtaata   | 1814 |
| Db | 1447 | CAGAACCCATGTCTTACTTTCAACGGATTAAGAAAGCATGTATGAGAA-----G          | 1497 |
| Oy | 1815 | aatgtlctatglttaagtaacagctgtaagatgacataagacagaaatlltacacagctactc | 1874 |
| Db | 1498 | CAGCGACAGATGTTAAATGCGACAGTAACGTACATTAACAGAAAGTTTACACAGCTACCTC   | 1557 |
| Oy | 1875 | tactataaazacaacaggaatgcattccttgacaattccgaaacacatatagtgaaatg     | 1933 |
| Db | 1558 | CATGTAAAAATGAACACAGAAAGGATTTCTGATGTTTTCTACACATTAATTAATTAATA     | 1617 |
| Oy | 1935 | acttaga---agaactaaggatagatcaactataatgtaacatacagagatctgacag      | 1991 |
| Db | 1618 | GAATGATTTAGAAACAGAGGGTGTGGACTCTGTAGGTGTGCCACGGTATGTGTACAG       | 1677 |
| Oy | 1992 | ttctcgtatcaglacagataatagtgatctcacaacaaacagaagtlcagatactgyc      | 2051 |
| Db | 1678 | TTTCTCATCTCAGTCCGAAATATAGTCTACCCCAACCAAAATTAACAGACAGATACTACAC   | 1737 |
| Oy | 2052 | ttttccagcaagagatctcaagaaaaggaccttcatccaatttactatgactctg         | 2111 |
| Db | 1738 | ATTTTTGACGAAAAGATTTTCAAGAAAAGACCTTATTTTGTGTGTCTATGATTTGTG       | 1797 |

|    |      |   |      |
|----|------|---|------|
| Qy | 2112 | gtctgataaacaataaacaagttccaagagaoactaacgtgttcaaggcaaaagctccatc   | 2171 |
| Db | 1798 | ATTCGGTGACATTAAATATCATCTCAAAAAGAACAGCTTAACAGTGAAGGCGGAACGGCATCT | 1857 |
| Oy | 2172 | atactccctccgaggaaccccatgtaatgtgacttcaagaatatgtatagttactctcg     | 2231 |
| Db | 1858 | GTGGTCTC---AGCAGCCCAATGAGTGTGACACGAGATATCAATATACCTTGCTTGTG      | 1914 |
| Oy | 2232 | gtaaaatacactcgaaaagtgaaaaataataattcaggaagaaatagaaaagaaactctgac  | 2291 |
| Db | 1915 | GTAGATCTCAGCGGGAAGGTAGACTGTTATTTAGCGCAGAG---AAAAAATGTGGACC      | 1971 |
| Oy | 2292 | caaatgcgtgaatttga---aaaagaactgaatttatcacagaagaagaacagaalt       | 2348 |
| Db | 1972 | CAATGCTGTAATTTGGATTAATAAAAAAGAACAGAGTTCTTCCATGTGTGAGACACAACTT   | 2031 |
| Oy | 2349 | gtatctaacccggtacagttcttctactagactgtttcagaactgtgaagagaatacgaat   | 2408 |
| Db | 2032 | GTGGGTCAACAGTACACTCTCTACTGGACCTTTTTCAAAACAAGCGGGAGAATTCAGATT    | 2091 |
| Oy | 2409 | tttgggtttcacagcttacacagaaaagaagtgaataatgaatgttttagatattgg       | 2468 |
| Db | 2092 | TTTTGGGCTTTACACAGCTACACTGAAAAACAGTGGCTATGTATGTATTAGATGTTTGG     | 2151 |
| Oy | 2469 | aagaagaaaattcagataactcgtltaacagcggttttccgcgtcccttcaacttcaat     | 2538 |
| Db | 2152 | AGGATGAAAAATTCAGATGATGTCGTCTGTCCAAACATTTTCTCCCTCCCTCAGCTTCAT    | 2211 |
| Oy | 2539 | ttactgacttttagaatttaaaaatgatcaacttttcagaagtgaataagatcatatct     | 2588 |
| Db | 2212 | TCAATGGATTTTACTATTATTAATGTGTTTTTCAGAAGTGACATTTTATCTTGGAAATTT    | 2271 |
| Oy | 2589 | tga-----aatttttaataatgaatgaaacttcttagaatllttttacag              | 2636 |
| Db | 2272 | TATATAATCATATATATAATTTTAGAGCTGCTTTTGGTTTTGGTTGTTTTCAGAC         | 2331 |
| Oy | 2637 | cttgttttacagaccacaatgtaataataataaataaataatttcgaatttctacaagt     | 2696 |
| Db | 2332 | CTTTGTTGCTGTGAACCATATGTAAATACATAAG--TGAATTTATTAATTTTACAGAAT     | 2389 |
| Oy | 2697 | tgaattaccgtltaagaanaaattacagaaataactgtgactgtgtctgttttaccattaa   | 2756 |
| Db | 2390 | TGAGTAGCTGTT--ACAATTAATACATAAATACTGTGATTATTAATTTACATACAAAAA     | 2447 |
| Oy | 2757 | aaaaaaaaa 2767  |      |
| Db | 2448 | AAAAAAAAAAAA 2458   |      |

|            |   |            |      |             |
|------------|---|------------|------|-------------|
| RESULT     | 7   |            |      |             |
| MMDBF4     |   |            |      |             |
| LOCUS      | MMDBF4  | 2309 bp    | mRNA | 11-NOV-1999 |
| DEFINITION | Mus musculus mRNA for DBP4-related protein, complete CDS.   |            |      |             |
| ACCESSION  | AJ003132  |            |      |             |
| VERSION    | AJ003132.1  | GI:6006562 |      |             |
| KEYWORDS   | DBP4-related protein.   |            |      |             |
| SOURCE     | house mouse.  |            |      |             |
| ORGANISM   | Mus musculus  |            |      |             |
| REFERENCE  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus |            |      |             |
| AUTHORS    | 1 (bases 1 to 2309)<br>Lepke,M., Putler,V., Stalb,C., Kneissi,M., Berger,C., Hoehn,K., Nanda,I., Schmidt,M. and Grunmt,F.           |            |      |             |
| TITLE      | Identification, characterization and chromosomal localization of the cognate human and murine DBP4 genes                            |            |      |             |
| JOURNAL    | Mol. Gen. Genet. 262 (2), 220-229 (1999)  |            |      |             |
| MEDLINE    | 99444905  |            |      |             |
| REFERENCE  | 2 (bases 1 to 2309)   |            |      |             |
| AUTHORS    | Lepke,M.  |            |      |             |
| TITLE      | Direct Submission   |            |      |             |
| JOURNAL    | Submitted (17-DEC-1997) Institute of Biochemistry, Am Hubland, Wuerzburg D-97074, Germany   |            |      |             |
| FEATURES   | Location/Qualifiers   |            |      |             |

source 1. .2309  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
gene 95. 2086  
/gene="MudBF4"  
95. 2086  
/gene="MudBF4"  
/function="Involved in initiation of DNA replication"  
/product="MudF4-related protein"  
/protein\_id="CAB56847.1"  
/db\_xref="GI:6006563"  
/translation="MLETMRHSKADLPQGTODRNRKPSLSTKADNRLEKSKV  
PLAKRFLIDSDSTICEKLOKDRKLGVEPEFLSKDIYFVSNNKAKYAOITLGRV  
SPVSPESKATLHETTSPHSDSSPKSDNRCTSGKILAKYKHDFIPANSLIS  
NALSKVYKLIHDIKITLIDKKALISATKSSASCKDAGKRAAGTOKPTGRLKRP  
FLKVEDNCRITREYLIQSLPCINLYLOKPCSPPIKESVVOKQNOYDPIYSDLY  
KCGEPVOLDJKEKRRKKECCLOKIDLETILSKHRNPNQSNQYDDESGHIL  
FDVEYGRDTPQKRRIRYSGSSVSANYLKNTAKERFLDPRNPDVGEYADLP  
KPNQYETQKPEERHGFASBPITYSSAGLKGDKPVMFASBPEDQBYADLP  
DSTPEHQVEGRNDQRPADPVSQSGVSHLSNLPOLADITOLSKDL  
QEKGFHVIIGHASDLVALNTSKQLMKARPCSPQEPHECDTEMNELPGKIQK  
VRMLLGQKANAEPSEALDKRTEYLPFAHBDRTGSPVQSLDLFQYSEKSEFLGT  
GYTNSGICVIDIWEENEESTLLSTFFSSPSTSAFVGF"

BASE COUNT 740 a 466 c 526 g 577 t  
ORIGIN

Query Match 44.5%; Score 1236; DB 10; Length 2309;  
Best Local Similarity 73.6%; Pred. No. 1.1e-224;  
Matches 1713; Conservative 0; Mismatches 570; Indels 45; Gaps 9;

QY 445 ggaagccgagcttggtgagcggtccgggtgagacattctccgagccagcatgtagtg 504  
Db 22 ggagccttggtggttgcgcgcgttccttcctgacgccccgggtggagccagcatgtagtg 81

QY 505 cgggagcagctgcagatgaaatcggagagatccagatgaaatcggagag 564  
Db 82 cctggagcagctgcagatgaaatcggagagatccagatgaaatcggagag 141

QY 565 tggagatcagatcgaatgaaatcgaatgaaatcgaatgaaatcgaatgaa 624  
Db 142 tggagatcagatcgaatgaaatcgaatgaaatcgaatgaaatcgaatgaa 201

QY 625 caggcagaagaatccaaatgaaatcgaatgaaatcgaatgaaatcgaatgaa 684  
Db 202 tggagatcagatcgaatgaaatcgaatgaaatcgaatgaaatcgaatgaa 261

QY 685 tctgcaacatcagatgaaatcgaatgaaatcgaatgaaatcgaatgaaatcga 744  
Db 262 tctgcaacatcagatgaaatcgaatgaaatcgaatgaaatcgaatgaaatcga 321

QY 745 agaatctcagatgaaatcgaatgaaatcgaatgaaatcgaatgaaatcgaatgaa 804  
Db 322 agaatctcagatgaaatcgaatgaaatcgaatgaaatcgaatgaaatcgaatgaa 381

QY 805 acaacctgggtgagatctccctgtaacagatcgaatgaaatcgaatgaaatcga 864  
Db 382 acaacctgggtgagatctccctgtaacagatcgaatgaaatcgaatgaaatcga 441

QY 865 caattcaacatcagatgaaatcgaatgaaatcgaatgaaatcgaatgaaatcga 924  
Db 442 caattcaacatcagatgaaatcgaatgaaatcgaatgaaatcgaatgaaatcga 501

QY 925 aagcagaagaatctagtgaaatgaaatcgaatgaaatcgaatgaaatcgaatgaa 984  
Db 502 aagcagaagaatctagtgaaatgaaatcgaatgaaatcgaatgaaatcgaatgaa 561

QY 985 tagatctatcaaatgagcttgtaagtgaaatgaaatcgaatgaaatcgaatgaaatgaa 1044  
Db 562 tagatctatcaaatgagcttgtaagtgaaatgaaatcgaatgaaatcgaatgaaatgaa 621

QY 1045 atactatcatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaa 1104

Db 622 ATACTCATTTGAAAGCAAGAAAGAAAGCACTGAGTGCATCAAGAAATCCAGCTCTTG 681  
QY 1105 aagagatgggggcaaaagatctgtagtgatgacacaaaacaaagacagagactgaa 1164  
Db 682 GAAGGATGCGGGGAAAAAGCTGCGCTGGAATACAAAGAGAAAGACAGGAGACTCAA 741  
QY 1165 aagacctcttgaaagtggaagatgagacaaacttatagaccatttatcttaagct 1224  
Db 742 GAAGCCTTTTAAAGGTTGAAGATGGAACCGTCTCAAGGCCATTTTACCTCAAGCT 801  
QY 1225 gacaaatgagctttaaataatctatctcagaagccctgagctcaatcttgatgaa 1284  
Db 802 CCCAGCTGCGGTGTATTAACCTCTTCTGAAAGCCCTGCACTCGTTGTGATATAG 861  
QY 1285 caagcactgtagtgcacaaagcaaacctcaggttaactaagaatccaaacagatggcga 1344  
Db 862 GAAGTATCCAGCGTCCAGAAAAACAAGCTCAGCCCAAGCTCAGAAATTAATATGACGCGCA 921  
QY 1345 taagtagtggaacatcgaatccagctcagtgaaagaaagaaagaaagaaatctg 1404  
Db 922 TAAAT---GTGGAACCCCAAGTTCAACCTCAGATTGAAGAAAAAGAAAGAAATATG 978  
QY 1405 tgaatgtgctgagcaagaatgagatgagatgagaaactcagctcagtgagcaacag 1464  
Db 979 TGAATGTGCTTACAGAAATACGAAGATCTCGAAACCTTCGAGTGAAGAAACAG 1038  
QY 1465 aaacttgacagagtaaccagatcaagtgatgagatgagatgagatgagatgagatg 1524  
Db 1039 GAATTTTGACACAGATACCAATCAAGTATGATGATGATGATGATGATGATGATGATG 1098  
QY 1525 tgaacttgagaaatgaaagagacacactaaagaaagaaagaaatcaatcagtg 1584  
Db 1099 CGATTTTGTGGAATATGAAAGAGACACCTCCTAAAGAAAGAAATGATGATGATGATG 1158  
QY 1585 atcccttcctcctgcttcgcagagtgctcgtgaaagaaagaaagaaagaaagaaag 1644  
Db 1159 GTCCCTTCTTCTGTTCTTCTGAAATGCTCTGAAAGAAAGAAAGAAAGAAAGAAAG 1218  
QY 1645 atgcaacatctcctcgaagaagatgcaagaagatgcaagaagatgcaagaagatgca 1704  
Db 1219 CTGGAGCCCAATTTCCAG 1278  
QY 1705 ttcctctataaagagacccagaaactgaaatgaaatgaaatgaaatgaaatgaaatgaa 1764  
Db 1279 CTGTCAGATGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1338  
QY 1765 ccccaacactcgaagaatgagatgagagagcttaatgagaaatgagaaatgagaaatg 1824  
Db 1339 AACCTATCTTCAAGCTGAGTGAAGATGGA-----CAGGAAACAGTGAAGCAT 1389  
QY 1825 gttaaagtcagctgaaagatgagatgagaaatgagaaatgagaaatgagaaatgagaa 1884  
Db 1390 gttcaatcgaagatgagctgagacccagagagagatgagacactgagctgagagag 1449  
QY 1885 caaagagatgcatctctgaaatctcgaacacacatgagtgaaatgagatgagaa 1944  
Db 1450 CACACCGGAAAC-----ACCAAGTGAACGAGGATGAAAGATGAC---GGAGA 1491  
QY 1945 actaaggtagatcaactataatgtaacatcagagcatctgtagatctgtagatctg 2004  
Db 1492 ACAGAGATGAGACCTGCTCAGAGCGTGTGCGAGTCAATGTGACAAAGTTTCACTCAG 1551  
QY 2005 tacaataatgtagatctcaacacaaagagagagagagagagagagagagagagagag 2064  
Db 1552 CACAAAGATATATACCCACACACAGTATAGCTCAGACATTACACACCTTTCAGCAAA 1611  
QY 2065 ggaactcagaagaagaagacttcatcaatattcaatcaatgagatctgagctcagaa 2124  
Db 1612 AATATGCAAGAAAGGCTTATGATGTTGATGATGATGATGATGATGATGATGATGATG 1671  
QY 2125 aaacagatcacaagaagacctaactgctcagaagaagagccatccatccatccatcc 2184



```

repeat_region 5950..6035
/rpt_family="MER1_type"
repeat_region 6360..6642
/rpt_family="Alu"
repeat_region 7010..7308
/rpt_family="Alu"
repeat_region 9806..10112
/rpt_family="Alu"
repeat_region 10303..10498
/rpt_family="L1"
repeat_region 10507..10807
/rpt_family="Alu"
repeat_region 10861..10978
/rpt_family="L1"
repeat_region 11533..11836
/rpt_family="Alu"
repeat_region 11959..12079
/rpt_family="Alu"
repeat_region 13992..14290
/rpt_family="Alu"
repeat_region 15571..16139
/rpt_family="L1"
repeat_region 16142..16438
/rpt_family="Alu"
repeat_region 16500..16779
/rpt_family="Alu"
repeat_region 16933..17077
/rpt_family="L1"
repeat_region 17850..18392
/rpt_family="L1"
repeat_region 18480..18769
/rpt_family="Alu"
repeat_region 18908..19366
/rpt_family="L2"
repeat_region 21782..21896
/rpt_family="L2"
repeat_region 22048..22156
/rpt_family="MER1_type"
repeat_region 24873..25250
/rpt_family="MALK"
repeat_region 26032..26309
/rpt_family="L2"
repeat_region 26342..26817
/rpt_family="MALK"
repeat_region 27147..27355
/rpt_family="Alu"
repeat_region 27381..29443
/rpt_family="L1"
repeat_region 29444..29561
/rpt_family="Alu"
repeat_region 31770..31909
/rpt_family="MIR"
repeat_region 32332..32720
/rpt_family="MALK"
repeat_region 32731..32953
/rpt_family="Alu"
repeat_region 33094..33327
/rpt_family="MER1_type"
repeat_region 33643..33941
/rpt_family="Alu"
repeat_region 34000..34088
/rpt_family="Alu"
repeat_region 34233..34602
/rpt_family="MALK"
repeat_region 38520..38721
/rpt_family="MIR"
repeat_region 38730..38849
/rpt_family="Alu"
repeat_region 40214..40605
/rpt_family="L2"
repeat_region 41565..41638
/rpt_family="L1"
repeat_region 42099..42347

```

```

repeat_region 42470..42784
/rpt_family="L1"
repeat_region 42787..43087
/rpt_family="Alu"
repeat_region 43588..43892
/rpt_family="Alu"
repeat_region 44036..44116
/rpt_family="L1"
repeat_region 44562..45028
/rpt_family="L1"
repeat_region 45191..51321
/rpt_family="L1"
repeat_region 52546..52617
/rpt_family="L2"
repeat_region 53530..53581
/rpt_family="L2"
repeat_region 54254..54761
/rpt_family="L1"
repeat_region 54759..55585
/rpt_family="L1"
repeat_region 55659..55857
/rpt_family="Alu"
repeat_region 58134..58298
/rpt_family="MIR"
repeat_region 59008..59181

Query Match 43.0% Score 1194.2; DB 9; Length 123331;
Best Local Similarity 99.7% Pred. No. 1,1e-216;
Matches 1196; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1563 aagaataaatacaagatgtagatcccttctctctggttctgcaagtgctgaaaga 1622
|||||
DB 94286 AAGAGTAAATACAGTGTGATCCCTTCTCTGTTTGCAGAGTCTGAAAAGA 94227

QY 1623 ctgaacaaagaagaagtcgaatcgacacatcttcagaaagatgccagaagatg 1662
|||||
DB 94226 CTGAACAAAGCAAAAGTGAATTCGAACATATTTCTCAGAAAAGTCCGGAAGTG 94167

QY 1683 atacacagtagaagaagaatcttcctgtataaagaagaccaggaactgaaaaagc 1742
|||||
DB 94166 ATACACAGTGAAGAGACCAATTTCTGTATTAAGAGACCCAGAACTGAAAAAGC 94107

QY 1743 tcccttatttagagagccatcccccacccctcaaatgaaatcgagaggttaagaga 1802
|||||
DB 94106 TCCCTATTATTTCAGACCCATCCCCACCCTTAATGAATGAGGGCTTAATGAGA 94047

QY 1803 aatagtaataaatgttcatttaagtaagcgtgaagatgacataaagaagaattta 1862
|||||
DB 94046 AATAGTAATAATGTTCATTTAAGTACAGCTGAAGATGACATTAAGCAGATTTTA 93987

QY 1863 ccaagctacccctacataaaacaacagaagatgcatcttgacatttcgaaccacat 1922
|||||
DB 93986 CCAAGCTACCCCTACATAAAACAACAGAAATGCAATTCGTGACATTTCCGAACACAT 93927

QY 1923 taagtgaataatgactagaagaactaaggtagatcactataaatgtaacataagcat 1982
|||||
DB 93926 TAAGTGAATAATGACTAGAAGAACTAAGGTGATCACTAATAATGATACAGGCAAT 93867

QY 1983 cgttcacgcttcctgattcagtagacagataatagtgatctcacaacaaagaagtcag 2042
|||||
DB 93866 CGTTCACGCTTCCTGATTTCAGTAGAGATTAATAGTGCATCAACCAAAACAGAGTCAG 93807

QY 2043 atacagtcctttcccaagaagaatcctcaagaagaagaacctcattcattatc 2102
|||||
DB 93806 ATACAGTCCTTTCCCAAGCAAGATCTCAAGAAAGACCTCATTCATTAATTTATTC 93747

QY 2103 atgatttcgctcgtataaataaacaagtcacaagcagacactcgttcagagcaag 2162
|||||
DB 93746 ATGATTTCGCTGTATAAATAAACAAGTTCACAAGACACACTAATGTTCAAGCAAG 93687

QY 2163 ctcacatcactccctccggaggaaacccaatgaaatgtaacttaagaataatgtagt 2222
|||||

```

|            |   |   |       |
|------------|---|---|-------|
| Db         | 93686   | CTCCATTCGCATACACCTCCTCCTCAGGACCAACCCATGAAATGAGCTTCAAGAAATATGCATATCT | 93627 |
| Qy         | 2223  | tacctctcgtgtaaaatalacatcgcgaagtgaaabaataltagacgcgaatagaaaaagaa      | 2282  |
| Db         | 93626   | TACCTTCGTGTAATAATACATGCAAAAGCAAAATAAATATATGAGCAAAATGAAAGANA         | 93567 |
| Qy         | 2283  | atctcgaaaccaaagctcgaatttgataaaagaactgaaatttatcaccaagaagaagaa        | 2342  |
| Db         | 93566   | ATCTCGAACCCAAAAGCTCGAATTTGTATAAAACACTGAAATTTATTACACMAAGAAAAACA      | 93507 |
| Qy         | 2343  | gaatttgagtcacccggtacagctccttacagactgtttcagactagtgaaagaat            | 2402  |
| Db         | 93506   | GAATTTGTAGTTCCACCGGTACAGTCTTTACTACACTCTTTCCAGACTAGTGAAGANAAT        | 93447 |
| Qy         | 2403  | cagaatttllgggtlltccaaagctacacagaaagaagtgatatalgcgaigtlttagata       | 2462  |
| Db         | 93446   | CAGAATTTTGGGTTCACAAGCTACACAGAAAAGAGGTATATGCAATGTTTAGATA             | 93387 |
| Qy         | 2463  | tttgggaagagaaaatctcagataatctgctttaaacaagtttctcgtccctccaact          | 2522  |
| Db         | 93386   | TTTGGGAAGAGAAAATTCACATATATCGTTTACACAGCGCTTTTCGTCCTCCCTCAACTT        | 93327 |
| Qy         | 2523  | ctacattcagcgcttttagaattttaaanaalgcatacttccagaagtgaagatca            | 2582  |
| Db         | 93326   | CTACATTTACTGCGCTTTTACAAATTTAAAAATGCCATCTTTTCAGAACTGATPAAGATCA       | 93267 |
| Qy         | 2583  | tattcttgaatttlltataaatafgatgtagaattccttagaatttlltaccagcttgt         | 2642  |
| Db         | 93266   | TATCTTGAAATTTTATATAAATATGATAGAAATTCCTTAGAATTTTACCAGCTTTGT           | 93207 |
| Qy         | 2643  | ttacagaccacaatgtaaatatltaaanaataatltgcaatttctacagaattgta            | 2702  |
| Db         | 93206   | TTACAGACCCTAAAGTAAATTTTAAAAATTAATTTTGCATTTTCTCACAGAAATGGAATA        | 93147 |
| Qy         | 2703  | ccctttaaagaaaattcacagaataacttggactgtgtctgttttaccatlaaaaaa           | 2761  |
| Db         | 93146   | CCTGTTAAAGAAAAATTCACAAATTAACCTGTGACTGCTGTGTTTACATTAATATATA          | 93088 |
| RESULT     | 9   |   |       |
| AC024947   |   |   |       |
| LOCUS      | AC024947  | 163031 bp   | DNA   |
| DEFINITION | Homo sapiens chromosome 01 clone RP11-200110, complete sequence.    |   |       |
| ACCESSION  | AC024947  |   |       |
| VERSION    | AC024947.4  | GI:9887599  |       |
| KEYWORDS   | HTG.  |   |       |
| SOURCE     | human.  |   |       |
| ORGANISM   | Homo sapiens  |   |       |
| REFERENCE  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;   |   |       |
| REFERENCE  | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.           |   |       |
| REFERENCE  | 1 (bases 1 to 163031)   |   |       |
| REFERENCE  | Smith,D.R.  |   |       |
| REFERENCE  | Genome Therapeutics Corporation Sequencing Center: Human Genome     |   |       |
| REFERENCE  | Sequence Data   |   |       |
| REFERENCE  | Unpublished   |   |       |
| REFERENCE  | 2 (bases 1 to 163031)   |   |       |
| REFERENCE  | Smith,D.R.  |   |       |
| REFERENCE  | Direct Submission   |   |       |
| REFERENCE  | Submitted (03-MAR-2000) Genome Therapeutics Corporation, 100 Beaver |   |       |
| REFERENCE  | Street, Waltham, MA 02453, USA                                      |   |       |
| REFERENCE  | 3 (bases 1 to 163031)   |   |       |
| REFERENCE  | Smith,D.R.  |   |       |
| REFERENCE  | Direct Submission   |   |       |
| REFERENCE  | Submitted (24-AUG-2000) Genome Therapeutics Corporation, 100 Beaver |   |       |
| REFERENCE  | Street, Waltham, MA 02453, USA                                      |   |       |
| REFERENCE  | On Aug 24, 2000 this sequence version replaced gi:8569066.          |   |       |
| COMMENT    |   |   |       |
| FEATURES   |   |   |       |
| FEATURES   | Location/Qualifiers   |   |       |
| FEATURES   | 1..163031   |   |       |
| FEATURES   | /organism="Homo sapiens"  |   |       |
| FEATURES   | /db_xref="taxon:9606"   |   |       |
| FEATURES   | /chromosome="01"  |   |       |
| FEATURES   | /clone="RP11-200110"  |   |       |
| FEATURES   | /clone_lib="RPC1-11"  |   |       |

[illegible]

|||||  
Db 36330 CTACATTTACGCTTTTAAATTAATAAATGCACTTTTACAGAGTATNAGAGTCA 36389  
2583 taattctgaatttttaataatgtaagaaattcttaagaattttttaccagcttgt 2642  
36390 TATTCCTGAAATTTTAAATGTATGGAATTTCTTAAGATTTTAAACGCTTGT 36449  
2643 ttacagaccacaatgtaatatataataataatcttgcaattttctacagaattgaata 2702  
36450 TTACAGACCCCAATGTAAATTAATAAATTAATTTCCATTTTCTACAGAAATTCATA 36509  
QY 2703 cctgtttaagaataattacagaataaactgtgactggtcctgttttacaataaaaaa 2761  
36510 CCTGTTAAAGAAATTAACAGATTAATTAATTTGACTGCTTGTTCATATATATATA 36568  
RESULT 10  
AC004958 163889 bp DNA HTG 12-JUN-1998  
LOCUS Homo sapiens clone D01096F22, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 6  
DEFINITION unordered pieces.  
AC004958  
VERSION AC004958.1 GI:3213039  
KEYWORDS HTG; HTGS\_PHASEL.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 163889)  
Waterson, R.H.  
JOURNAL The sequence of Homo sapiens clone  
Unpublished  
2 (bases 1 to 163889)  
Waterson, R.H.  
REFERENCE Direct Submission  
Submitted (12-JUN-1998) Genome Sequencing Center, Washington  
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108 USA  
COMMENT \* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 6 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 1462: contig of 1462 bp in length  
\* 1463 1480: gap of unknown length  
\* 1481 3064: contig of 1584 bp in length  
\* 3065 3082: gap of unknown length  
\* 3083 4710: contig of 1628 bp in length  
\* 4711 4728: gap of unknown length  
\* 4729 6530: contig of 1802 bp in length  
\* 6531 6547: gap of unknown length  
\* 6548 16429: contig of 9882 bp in length  
\* 16430 16446: gap of unknown length  
\* 16447 163889: contig of 147443 bp in length.  
Location/Qualifiers  
1..163889  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="D01096F22"  
BASE COUNT 49499 a 30566 c 31570 g 52166 t 88 others  
ORIGIN  
Query Match 43.0%. Score 1194.2. DB 2. Length 163889;  
Best Local Similarity 99.7%. Pred. NO. 1.1e-216.  
Matches 1196; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1563 aaagaataaatacagtgctgaccccttcctcgtttctcgaagtgctcctgaagaaga 1622  
|||||  
Db 74946 AAGCAATTAATACAGTGTGATCCCTTCTCCTCTTCTGCAAGTGCTCGAAGAAAGA 75005

QY 1623 ctgaacaaaggagaaagtggaattgcacatacttcagaagaatgtccagaagaatg 1682  
Db 75006 CTGACCAAAAGGAAAAAGTGGAAATTCACATATTTCTCGAAGAAATTCGACGAGAAATG 75065  
QY 1683 atcaacaatgtgaagagagaagaattccctgtatagaagaccagaagaactgaaaaaagc 1742  
Db 75066 ATACACAGTGAAGAGACAGAAATTCCTGTATTAAGAGACCAGGAATCTAAAAAAGC 75125  
QY 1743 tctgtttattcagaagccatcccccacccctcaatgaattgagaggttataaga 1802  
Db 75126 TCTGTTATTTTCAGAGCCATCCGCCCTTCATTAATTAATGAGAGCGCTTATATAGA 75185  
QY 1803 aaatgaataaattgttccatglttaagtaacagctgaagaatgaatgaagaagaatttta 1862  
Db 75186 AAATGACTAATAAATGTTCCATGTTAAGTACAGCTGAAGATGACATTAAGACAGAAATTTTA 75245  
QY 1863 cacaagctacctacataaacaagaagaatgtcattcttgacatttcogaacacat 1922  
Db 75246 CACAGTACTCTTACATATAAAACAACAGAAATGCAATCTTGACATTTCCGAAACACAT 75305  
QY 1923 taagtgaataatgactagaagaagactaagggtagaatcactataatgtatacacaaggcat 1982  
Db 75306 TAACTGAAATGACTTAGAAGAACCTTAAGGTAGATCACTATTAATGTAACTACAGGCAT 75365  
QY 1983 ctgtacaatgttctgaattctcagtaagaataatagtgatctcacaagaagaagatcag 2042  
Db 75366 CTGTACATGTTTCTGATTTCACTACAGATTAATGATGATCTCACCAAAACAGAAATCTAG 75425  
QY 2043 atactgtcttttccagaagaagatctcagaagaagaagacttcataatattactc 2102  
Db 75426 ATACGTGCTTTTCCAGAAAGATCTCAAGGAAAGACCTTCATTAATTTACTC 75485  
QY 2103 atgattctgtctgtataacaataaacagttcacaagaagcccaactgttcaagaagaag 2162  
Db 75486 ATGATTTGTGTCGTGATTAACAAATTAACACTTCAAGAGCACTTACTGTTCCAGCAAAAG 75545  
QY 2163 ctcatctccatctcctctcttgagaagaccgaatgaatgtgactcagaatatgtatgtc 2222  
Db 75546 CTCATTTCCATCTCTCTCTGAGAAACCAATGAATGTGACTTCAAGAAATATGATGATGAT 75605  
QY 2223 tacctctgtgtaaaatacacacgaagaatgaataataatagaagcgaatagaagaaga 2282  
Db 75606 TACCTTCGTGTAATAATATACATAAAGTGAATAATATATGAGACCAATTAAGAAAGAA 75665  
QY 2283 atctggaaaccaatgtcgaatttgaataaagaactgaatttattacacaagaagaaga 2342  
Db 75666 ATCTGGAACCAAAATGCTGAATTTGATTAAGACAGTAATTTATACAAAGAAAGAA 75725  
QY 2343 gaatttgaattcaacggtaacagttcttactagaactgtttcagaactagtagaagaat 2402  
Db 75726 GAATTTGTATGTTCCGCTGACACTCTTACTAGACTGTTTCAAGCTAAGTAAGAAAT 75785  
QY 2403 cagaatttttgggtttcacaagcttacaagaagaagaagtgatatacaatgtttaagta 2462  
Db 75786 CAGAAATTTTGGGTTTACAGCTACAGAAAGAGTGTATATGCAATGTTTAAAGTA 75845  
QY 2463 ttctgggaaggagaatcaatcaagataatcgttacaaggttttctcgtcccttaact 2522  
Db 75846 TTTGGGAAGAGAAATTCAGATATCTGTTAACAGCGTTTCTGCTCCCTTCAACTT 75905  
QY 2523 ctacattactgtgcttttgaatttaaaatgaatcattcttcaagaatgtataagaatca 2582  
Db 75906 CTACATTTACGCTTTTGAATTTTAAATTAATAAATGCAATCTTTTCAAGAAATGATTAAGATCA 75965  
QY 2583 taattctgaatttttaataatgtaagaaattcttaagaatttttttaccagcttgt 2642  
Db 75966 TATTCCTGAAATTTTAAATATGATGAGAAATCTTAAAGATTTTAAACAGCTTGT 76025  
QY 2643 ttacagaccacaatgtaatatataataataatcttgcaattttctacagaattgaata 2702  
Db 76026 TTACAGACCCCAATGTAAATTAATAAATTAATTTTCAATTTTCTACAGAAATTAATA 76085



QY 2703 cctgtttaaagaataatcacagaataactgtgactgctgttcttcttacaataaaaaa 2761  
|||||  
Db 76086 CCTGTTAAAGAAAAATTAACGAATTAACCTTGACGTGCTTCTTTTACATTATATATA 76144

RESULT 11  
AK022969 2276 bp mRNA PRI 29-SEP-2000  
LOCUS AK022969  
DEFINITION Homo sapiens CDNA FLJ12907 fls, clone NT2RP2004396, highly similar  
to Homo sapiens mRNA for activator of S phase Kinase.  
ACCESSION AK022969  
VERSION AK022969.1 GI:10434668  
KEYWORDS oligo cloning; fls (full insert sequence).  
SOURCE Homo sapiens teratocarcinoma cell line:NT2 CDNA to mRNA,  
clone\_11b:NT2RP2 clone:NT2RP2004396.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (sites)  
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,  
Nishikawa,T., Nagai,K., Sugano,S., Aotsuka,S., Yoshikawa,Y.,  
Matsunawa,H., Ishii,S., Kawai,T., Saito,K., Yamamoto,U.,  
Wakamatsu,A., Nakamura,Y., Nagahari,K., Masudo,Y. and Sasaki,N.  
TITLE MEDO human cDNA sequencing project  
JOURNAL Unpublished (2000)  
REFERENCE 2 (bases 1 to 2276)  
AUTHORS Isogai,T. and Otsuki,T.  
TITLE Direct Submission  
JOURNAL Submitted (23-AUG-2000) to the DDBJ/EMBL/GenBank databases. Takao  
Isogai, Helix Research Institute, Genomics Laboratory: 1532-3 Yana,  
Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hrl.co.jp,  
Tel:81-438-52-3951, Fax:81-438-52-3952)  
COMMENT MEDO human cDNA sequencing project supported by Ministry of  
International Trade and Industry of Japan; CDNA full insert  
sequencing; Research Association for Biotechnology; CDNA library  
construction, 5'- 3'-end one pass sequencing and clone selection;  
Helix Research Institute (supported by Japan Key Technology Center  
etc.) and Department of Virology, Institute of Medical Science,  
University of Tokyo.  
FEATURES  
Source  
1..2276  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_line="NT2"  
/cell\_type="teratocarcinoma"  
/clone="NT2RP2004396"  
/clone\_11b="NT2RP2"  
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal  
precursor cells after 2-weeks retinoic acid (RA)  
induction."  
BASE COUNT 793 a 381 c 421 g 681 t  
ORIGIN

Query Match 39.6%; Score 1101.6; DB 9; Length 2276;  
Best Local Similarity 99.6%; Pred No. 3.5e-199;  
Matches 1104; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1654 tatttctcagaagaatctgcagaagaatgatacaacagttgaagagacagaatttcctgta 1713  
|||||  
Db 1 TATTTCTCAGAAAGATTGCCAGAGATGATACACAGTAAAGAGAGAGATTTCCGTGA 60

QY 1714 taaagagaccaggaactcgaataaagctcgttatttctcagagccatcccccacc 1773  
|||||  
Db 61 TAAAGGACCCAGGAAATCGAAAAAAGCTCCGTATTATTCAGAGCCATCCCCACCC 120

QY 1774 ttcaatgagttgagagggcttaataagaataaagttaataaattgttcacatgttaagtac 1833  
|||||  
Db 121 TTCAAATGAAATGAGGGCTTATATGAGAAATGAGTAATAAATGTTCCATGTTAAAGTAC 180

QY 1834 agctgaagatgacataagaacagatttcaacagctacacctcatataaacaacagga 1893  
|||||  
Db 181 AGCTGAAGATGACATAGACAGCAATTTTACACAGTACCTCTACATTAACAAACAGGA 240

QY 1894 atgcattctgacatttcggaacacacaaatgaagtgaanaatgacttagaagaagagct 1953  
|||||  
Db 241 ATGCATCTTGACATTTCCACACACACATTAAGTGAAGAAAGACTTGAACAACTAAAGGCT 300

QY 1954 agatcacataaattgtacatacagagcatctgacatgcttcctgattcagttacagataa 2013  
|||||  
Db 301 AGATCACTATAAATGTATACATACAGAGCATCTGTACATGTTTTCATGATTTCACTACAGTAA 360

QY 2014 tagtggatctcaacccaacagaagtcagatgactgtcttctccagcaagagctcaa 2073  
|||||  
Db 361 TAGTGATCTCAACCAACAAACAGAAAGTACAGTACTGTCTTTTCCAGCAAGATCTCAA 420

QY 2074 ggaagaagaccttcattcaatatttctactcagttctgctgcgcgagatacaataaagcttc 2133  
|||||  
Db 421 GGAAAGGACCTTCATTCAATATTACTCATGATTCGTGCTGATTAACAAATTAACACTTC 480

QY 2134 acaagagcacctactgttcaagcaaaagctccattccatacctccctcctgaagaaacca 2193  
|||||  
Db 481 ACAAGAGCACCTTAACGTTCAGGCAAAAGCTCCATTCATACCTCTCTGAGCAACCCAA 540

QY 2194 tgaatgtgactccaagaataatgatatgattacacctctgtaaaaaatatacgcgaagtga 2253  
|||||  
Db 541 TGAATGTGACTTCAAGCAATATGATAGTTTACCCTCTGTGAATAATCATCGAAAGTGAA 600

QY 2254 aataatattagagcgaataatgaagaanaatctggaacccaatgctgaaatttgataaag 2313  
|||||  
Db 601 AATTAATATTAGAGAGCAAAATGAAAAAGAAATCTGGAACCAATGCTGAATTTGATTAAG 660

QY 2314 aactgaaattattacacaagaagaanaacagaattgtgactcaccgttacagtttact 2373  
|||||  
Db 661 AACTGAATTTATACACAAAGAAAAACAGAAATTTGATGTTACACGGTTACAGTCTTACT 720

QY 2374 agactgttccaagactggaagaagaatcaagaatttgggttccaagactacacaga 2433  
|||||  
Db 721 AGACTGTTTCAGACCTGTGTGAGAGAAATCAGAAATTTTGGGTTTCAACAGCTACACAGA 780

QY 2434 aaagagagatattgacatttttagattttggaagaggaagaataatcacatattcgt 2493  
|||||  
Db 781 AAAGAGGTATATGCAATATTTAGATTTGGGAAGAGGAAAAATTCAGATTAATCTGTT 840

QY 2494 aacagcgtttctcgcgccctccaacttccattacattactgacctttagaatttaaaaa 2553  
|||||  
Db 841 AACAGCGTTTTTCTCGGCCCTTCACTTCTACATTACAGGCTTTTGAATTTAAAAAA 900

QY 2554 tgcatacttttcagaagtgataagagatcatalctctgnaatttttaataatgatatga 2613  
|||||  
Db 901 TGCATACTTTTCAGAAAGTAAAGATCATATTCTTGAATTTTATTAATAATGTATGGA 960

QY 2614 aattcttaggaattttttaccagcttggtttacagaccaccaatgtaatatlaaataa 2673  
|||||  
Db 961 AATTCCTTAGGATTTTTTTTACCAGCTTTGTTTACACAGCCCAAAATGTAATTTAAAAATTA 1020

QY 2674 atatttgaatttctcagaatgaattgaattcgttgaagaanaatcagaataactg 2733  
|||||  
Db 1021 ATATTTCGAATTTTCTACAGAAATTTGAATTTACCTGTTAAAGAAAAATTTACAGAAATAACTTG 1080

QY 2734 tgaactgtctgttttactaataaaaaa 2761  
|||||  
Db 1081 TGACTGCTCTTGTGTTTACATTATATATA 1108

RESULT 12  
AC091341/c  
LOCUS AC091341  
DEFINITION Rattus norvegicus clone CH230-1C19, WORKING DRAFT SEQUENCE, 35  
unordered pieces.  
ACCESSION AC091341  
VERSION AC091341.1 GI:13661907  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;







```
repeat_region 6768..7030 /rpt_family="MALR"
repeat_region 7272..7610 /rpt_family="Retroviral"
repeat_region 7403..7628 /rpt_family="MER4-group"
repeat_region 7660..7898 /rpt_family="Retroviral"
repeat_region 8013..8755 /rpt_family="L1"
repeat_region 8777..9094 /rpt_family="Alu"
repeat_region 9389..10058 /rpt_family="MER1-g"
repeat_region 10069..10231 /rpt_family="L1"
repeat_region 10626..11262 /rpt_family="L1"
repeat_region 11263..11827 /rpt_family="Retroviral"
repeat_region 11828..12240 /rpt_family="L1"
repeat_region 12241..12532 /rpt_family="Alu"
repeat_region 12533..12548 /rpt_family="L1"
repeat_region 12549..12849 /rpt_family="Alu"
repeat_region 12850..13189 /rpt_family="L1"
repeat_region 13205..13372 /rpt_family="Alu"
repeat_region 13376..13881 /rpt_family="L1"
repeat_region 13882..14120 /rpt_family="L1"
repeat_region 14084..14283 /rpt_family="L1"
repeat_region 14286..14760 /rpt_family="L1"
repeat_region 14765..15133 /rpt_family="MALR"
repeat_region 15316..15669 /rpt_family="MALR"
repeat_region 17202..17287 /rpt_family="MALR"
repeat_region 17818..18025 /rpt_family="L1"
repeat_region 18030..18105 /rpt_family="L1"
repeat_region 18115..18424 /rpt_family="Alu"
repeat_region 18791..18835 /rpt_family="(CAT)n"
repeat_region 18888..19165 /rpt_family="Alu"
repeat_region 19778..20821 /rpt_family="Retroviral"
repeat_region 20875..20991 /rpt_family="MIR"
repeat_region 21983..42089 /gene="WUGSC:H_RG060N22.1"
CDS join<21983..72046,24028..24134,28390..28457,30361..30485,
31216..31389,34475..34584,36624..36705,41528..41608,
41977..42089)
/gene="WUGSC:H_RG060N22.1"
/note="mitochondrial carrier protein-like; similar to
009461 (PID:g2497990); H_RG060N22.1"
/codon_start=1
/evidence-not_experimental
/protein_id="AAC82534.1"
/db_xref="GI:2588623"
/translation="IVPLDVVKIRLQANNPLPKGKCFVYSNGLMDHLCVCEBGNK
```

```
LMYKKPNFOGLDAFFKILRNEGKLSLGLPPTLVAAVPATVYFTCYDLSALLR
SKLGENERICPLVAGIVARFQWVYVLSLEILRTMOGSKFSYLFHFFVSKVSEGD
MISLMRGAPVIVLRDVPFSAWVYVLETLKILCKSLYPTPIINTSGALSSEFA
AVATLPEFVYVTKORTOLMTYESHKSISPLHNSWTINKNIYAKNGFSGLSGLPRL
IKIAPACAIMISTYEFKAFPOKONVROOY"
repeat_region 23159..23457 /rpt_family="Alu"
repeat_region 23507..23784 /rpt_family="Alu"
repeat_region 24328..25740 /rpt_family="L1"
repeat_region 25741..26036 /rpt_family="Alu"
repeat_region 26037..26232 /rpt_family="L1"

Query Match 19.9%; Score 554.6; DB 9; Length 136823;
Best Local Similarity 99.3%; Pred. No. 2,7e+95;
Matches 557; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 cggacagagctctctgagctcgcccaagacctgaagcggcgagaccgagccgggtct 64
DB 2118 CGCGCCCAACTCTCTGAGCGCTGCGCCAAAGACTGAAGCGGCGGACCGAGCCCGGGTCT 2059
QY 65 ggactgagagagcaacggaatgagcggggtgagcgcggaacacactgcaaggcc 124
DB 2058 GAGACTGAGAGAGCAAGGAAATGAGAGCGGGGTAGAGCGCAACAACCTGCAGAGGCC 1999
QY 125 aagacgaagcgagaaagacgagcgctgagggcgagggcgcgagcgcgagaaagca 184
DB 1998 AGAGCGAGGCGGAGAAAGAGCGGCGGCTGAGGGGCGGCGCGAGCGGAAGGCA 1939
QY 185 ggcacgaaggcgagcgagcgagggcgagcgcgctgagcgagcgagggcgagcg 244
DB 1938 GGCACGAGGCGGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1879
QY 245 cgtatcgagcgcgagcgcgctgagcgcttccaatctcaacgcgcgagccactcg 304
DB 1878 CGATCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1819
QY 305 ttgtgcttgcgctctcctcgcgcttggagcgagatccggcccgaaaccca 364
DB 1818 TTGTGCTTGGGCGCTTCCCTCGCGGCGGCTTGAGACCGGATCCGGCCCGGAACCGCA 1759
QY 365 cctgcagacgagctactactgctagagcgctagctgaggaagagagagcgagc 424
DB 1758 CCGCAGACGCGGCTACTGCTAGCGCTGAGGCGGCTGCGGGAAGAGAGAGCGGCG 1699
QY 425 cgtcctgtaaaagcgcgaggagcggtgcttgcgagctgcccgggtcgagacttct 484
DB 1698 CGTCTGTCAACAGCGCGGCGGGAAGCGGTGCTTCCGCGCTGCGCGTGGACACTTCT 1639
QY 485 cggaccacgcatgtagtgcgcggcgagctgcatgaactccggaagcatalgagatccac 544
DB 1638 CCGGACCCAGCATGTAGTGCAGCGGCGGACTGCCATGAACCCGAGCAATGAGATCCAC 1579
QY 545 agtaagagacattccagaggt 565
DB 1578 AGTAAAGACATTTCAGGGT 1558

RESULT 14
AX070167 685 bp DNA PAT 25-JAN-2001
LOCUS AX070167
DEFINITION Sequence 639 from Patent WO0102568.
ACCESSION AX070167
VERSION AX070167.1 GI:12579952
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 685)
```

AUTHORS Williams, L.T., Escobedo, J., Innis, M.A., Garcia, P.D., Klinger, J., Kasam, A., Reinhard, C., Randazzo, F., Kennedy, G.C., Pot, D., Lamson, G., Drmanac, R., Crkenjakov, R., Drmanac, S., Dickson, M., Labat, I., Leshkowitz, D., Kita, D., Garcia, V. and Strache-Crain, B.

TITLE Human genes and gene expression products

JOURNAL Patent: WO 0102568-A 639 11-JAN-2001; CHIRON CORPORATION (US); HYSEQ, INC. (US)

FEATURES

source 1. 685

/organism="Homo sapiens"

/db\_xref="taxon:9606"

misc\_feature 1. 685

/note="n = A,T,C or G"

BASE COUNT 238 a 127 c 125 g 187 t 8 others

ORIGIN

Query Match 18.5%; Score 513.8; DB 6; Length 685;

Best Local Similarity 94.9%; Pred. No. 1,2e-87;

Matches 360; Conservative 0; Mismatches 24; Indels 6; Gaps 3;

QY 558 tccagggtggaatccaaagtcacaaatgaaaaaacagaccatctctgaatctctgaaaa 617

DB 1 TCAGAGGTGGAATCCAAAGTCAAAATGAAAAAACACCATCTCTGAAATCTCTGAAA 60

QY 618 ctgatacagagcagaaaaatccaaatgtaagccacttggggaagaatatttaccttg 677

DB 61 CTGATTAACAGGCCAGAAAAATCCAAATGTAAAGCCACTTGGGGAAAAAGTATTTCCTTG 120

QY 678 actacatctctgacccatctctgaaaaactcaaaagacataagagatctgggagggc 737

DB 121 ACTTACTCTCTGCAACCAATCTGAAAAACTTCAAAAAGACATTAAGATCTGGAGAGGC 180

QY 738 gacgtgaaagatctcagcaagaagatacagttatcttatttcaataagaagaacct 797

DB 181 CAGTGAAGAAATTTTCAGCAAAAGATATCATTTATCTTCAATTAAGAAAGAACCTA 240

QY 798 aattgcaaaactctggtgcaatctctcttaccagtcacagatccatcatctact 857

DB 241 AATTGCAAACTTGGTGGCAATTTCTCTGTACCAAGTCACAAATCTCATTAATCTG 300

QY 858 cagaaccacttaccatccagcagcagatgaggaagttcatttaagtcacacagacag 917

DB 301 CAGMAACCACTTACCTCATCCACGACATGAGTGAAGTCAATTAAATGACCAACAG 360

QY 918 tgggttaagcaggaagaatattatagttgaaaaagcatcaagagacagatttcttc 977

DB 361 TGTGTTAAACCAAGCAAAATTTAGTTGAAAAAGCTATCAAGGACATGATTTTATTC 420

QY 978 ctcaaatgatatatatcaaatgcttgcattgagggagtaaaatcttcataatgag 1037

DB 421 CTTCAAATATATATATATCAATGCTTGTATGAGNAGTANNAATCTTCATATTTATG 480

QY 1038 acattatattctcatctgacacaaagaagaagtgatttatttactcaagaagaacagta 1097

DB 481 ACATTATATCTCATCTGTAACANAGAAAAAGAGNT--GATTACTACAGAA--TCANGA 536

QY 1098 cttaagaagaagatggggcagaagaagtgtagtggtagtggtagcacaacaaaca 1147

DB 537 CTTCATTAANNAATGGCG--CAAAAAAGTTCTAGTGTGTCCTAATACCA 584

RESULT 15

AC074175 177899 bp DNA HTG 15-JUL-2000

LOCUS Mus musculus chromosome 3 clone RP23-348B11 strain C57BL/6J,

DEFINITION WORKING DRAFT SEQUENCE, 15 unordered pieces.

ACCESSION AC074175.1 GI:9211286

VERSION AC074175.1

KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc.mouse@nih.gov

Center project name: vc

Center clone name: 348B11

----- Summary statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 171084 bases at least Q40

Consensus quality: 172951 bases at least Q30

Consensus quality: 174122 bases at least Q20

Insert size: 187000; agarose-fp

Insert size: 202000; pulse-field-gel

Insert size: 176499; sum-of-contigs

Quality coverage: 4.87x in Q20 bases; agarose-fp

Quality coverage: 4.50x in Q20 bases; pulse-field-gel

Quality coverage: 5.16x in Q20 bases; sum-of-contigs

-----

NOTE: This is a 'working draft' sequence. It currently consists of 15 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as \* runs of N, but the exact sizes of the gaps are unknown. \* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 2463: contig of 2463 bp in length

2464 2563: gap of unknown length

2564 4767: contig of 2204 bp in length

4768 4867: gap of unknown length

4868 8397: contig of 3530 bp in length

8398 8497: gap of unknown length

8498 13167: contig of 4670 bp in length

13168 13267: gap of unknown length

13268 17890: contig of 4523 bp in length

17891 23556: gap of unknown length

23557 23656: contig of 3666 bp in length

23657 32999: gap of unknown length

32999 33099: contig of 9343 bp in length

33000 44567: gap of unknown length

44568 44667: contig of 11468 bp in length

44669 56459: gap of unknown length

56460 72129: contig of 11792 bp in length

72130 72229: gap of unknown length

72230 87873: contig of 15570 bp in length

87874 87973: gap of unknown length

87974 108836: contig of 15644 bp in length

108837 108936: gap of unknown length

108937 130736: contig of 20863 bp in length

130737 130736: gap of unknown length

130738 151003: contig of 21700 bp in length

151003 20267: gap of unknown length

20267 20267: contig of 20267 bp in length



Fri Dec 28 08:22:45 2001

us-09-830-647-3.rge

Page 24

